

UNIVERSIDADE DE BRASÍLIA

ANDRÉ LUIZ DE FARIA LEITE

**INCONGRUÊNCIAS EM FALA LIDA: ESTUDO DE CORPUS DE
PRONUNCIAMENTOS PARLAMENTARES**

BRASÍLIA (DF)

2021

ANDRÉ LUIZ DE FARIA LEITE

Incongruências em fala lida: estudo de corpus de pronunciamentos parlamentares

Dissertação submetida ao Programa de Pós-Graduação em Linguística do Departamento de Linguística, Português e Línguas Clássicas do Instituto de Letras da Universidade de Brasília como requisito parcial para obter o título de Mestre em Linguística

Orientadora: Prof^a Dra. Aveliny Mantovan Lima

Brasília (DF), Brasil

2021

Ficha catalográfica

FF224i Faria Leite, André Luiz de Incongruências em fala
lida: estudo de corpus de pronunciamentos parlamentares /
André Luiz de Faria Leite; orientador Aveliny Mantovan
Lima. -- Brasília, 2021. 297 p.

Dissertação (Mestrado - Mestrado em Linguística) -
Universidade de Brasília, 2021.

1. leitura. 2. fluência. 3. prosódia. 4. psicolinguística.
5. memória de trabalho. I. Mantovan Lima, Aveliny, orient.
II. Título.

AGRADECIMENTOS

À professora Aveliny Mantovan Lima, que teve a bondade de me acolher como orientando e a coragem de me deixar percorrer livremente as arriscadas veredas da linguística.

Aos membros da banca examinadora: professora Camila Leite Tavares, que, de certa forma, coorientou o trabalho com suas observações na fase de qualificação; e professor Dionei Moreira Gomes, que enfileirou críticas bastante pertinentes a determinadas hipóteses.

Aos professores e professoras do Programa de Pós-Graduação em linguística da UnB — Walkiria Neiva Praça, Rozana Reigota Naves, Paulo Medeiros Júnior, Marcus Vinicius da Silva Lunguinho e Dionei Moreira Gomes — que, cada qual à sua maneira, mas todos com muita empatia, benevolência e comiseração, contribuíram para a designorância linguística do autor.

Aos voluntários, especialmente à Mari e ao Ian, que, em uma longa noite de verão em Garopaba, doaram tempo e paciência para validar as incongruências de fronteiras prosódicas.

E, especialmente, à Daniela, ao Daniel e ao Davi, que testemunharam a laboriosa construção deste trabalho, por vezes com surpresa ou curiosidade, mas sempre com bom humor e compreensão.

Hoje eu vim, minha nega
Sem saber nada da vida
Querendo aprender contigo
A forma de se viver
As coisas estão no mundo
Só que eu preciso aprender
As coisas estão no mundo
Só que eu preciso aprender

(Paulinho da Viola, 1968)

RESUMO

Contexto: A literatura linguística não tem produzido muitos dados sobre a fala lida (leitura em voz alta) em indivíduos adultos e escolarizados. A produção de dados desse tipo pode caracterizar melhor o desempenho dos falantes nessa tarefa e evidenciar certos mecanismos de processamento da linguagem que podem ter caráter universal.

Objetivos: a) identificar incongruências involuntárias entre textos escritos e sua leitura em voz alta (erros de leitura); b) correlacionar as incongruências identificadas com aspectos morfosintáticos, semânticos e discursivos dos textos; e c) reconhecer, na análise das correlações, possíveis mecanismos fundamentais de processamento da linguagem.

Métodos: Identificaram-se os erros de leitura produzidos em um corpus composto por 5 horas de vídeos de falas lidas, com um tempo médio de fala de 8 minutos por falante, incluindo 39 indivíduos, com idade média de 63 anos, de sexo predominantemente masculino e nível de escolaridade majoritariamente superior. Os erros foram correlacionados com variáveis demográficas e linguísticas, e as correlações foram submetidas a análises estatísticas.

Resultados: As falas foram segmentadas em 1429 períodos. A maioria dos períodos (854 períodos) continham algum tipo de incongruência. As incongruências mais comuns foram trocas de palavras (16%), desvios de ortoépia (15%), desvios de concordância (14%), pausas (13%) e fronteiras prosódicas incongruentes (20%). Houve correlação positiva entre incongruências e: a) sexo masculino (mulheres cometeram menos erros, p -value = 0,0041), e b) extensão dos períodos. Não houve correlação estatisticamente significativa entre incongruências e: a) idade, e b) taxa de elocução (velocidade de fala). No caso específico de incongruências de fronteiras prosódicas (segmentação prosódica dos enunciados), observou-se correlação: a) positiva entre incongruência e continuidade temática entre períodos (o tópico do período prosodicamente incongruente havia sido mencionado no período anterior, p -value 0,0006345) e b) negativa entre incongruência e sujeito em primeira pessoa (p -value = 0,0002584). Observou-se também forte correlação positiva entre incongruências de fronteiras prosódicas e um determinado critério de segmentação dos períodos que considera que determinadas palavras poderiam ser omitidas sem comprometer a composição semântica e gramatical mínima dos enunciados e de seus segmentos (p -value < 2,2e-16). No caso das trocas de palavras, observou-se que: a) 80% das palavras trocadas são parônimos (palavra falada é semelhante à palavra escrita), e b) as palavras faladas costumam ser mais frequentes na língua do que as palavras escritas. Com relação ao fenômeno de retratação (autocorreção do erro de leitura), constatou-se que: a) houve retratação em 26% dos eventos de incongruência segmentar e b) não houve retratação em virtualmente nenhum dos eventos de incongruência suprasegmentar (prosódica). Além disso, as retratações de trocas de palavra e desvios de ortoépia tendem a obedecer a um curioso paradigma em que o substantivo retratado é acompanhado pelo artigo ou preposição antecedente.

Conclusões: Erros de leitura são bastante comuns entre falantes adultos escolarizados, e associam-se a certos padrões linguísticos que podem refletir mecanismos cognitivos fundamentais de processamento da linguagem. A esse respeito, elaboram-se hipóteses explicativas e apresentam-se algumas conjecturas, dando-se especial ênfase ao conceito de memória de trabalho. Sugerem-se, também, estratégias de produção textual que poderiam facilitar a leitura em voz alta.

Palavras-chave: leitura; fluência; prosódia; psicolinguística; memória de trabalho

ABSTRACT

Background: *linguistics has not gathered enough data about oral reading fluency in adult, educated speakers. Such data may shed some light on the subject and reveal essential mechanisms of sentence processing.*

Objectives: *a) to determine oral reading errors prevalence among adult readers, b) to correlate the errors with linguistic aspects of the written texts, and c) to identify, in the correlations, potential sentence processing mechanisms.*

Methods: *We identified oral reading errors in a corpus comprising five hours of videos of oral readings made by 39 adult Brazilian Portuguese native speakers, primarily college-educated males, with a mean age of 63 years old. Correlations with linguistic and demographic variables were statistically analyzed.*

Results: *Oral readings were segmented into 1429 sentences. Most of the sentences (854) had some type of incongruence with the written texts. The most frequent incongruences were word substitutions (16%), cacoeopies (15%), agreement errors (14%), pauses (13%), and prosodic boundary incongruences (20%). There was a positive correlation between errors and a) male gender (p -value = 0.0041) and b) sentence length, and no correlation with c) age and d) speech rate. Prosodic boundary disfluencies correlated positively with thematic continuity (disfluent sentence topic has been mentioned in the previous sentence, p -value = 0,0006345) and negatively with first-person subjects (p -value = 0.0002584). We also found a strong correlation between prosodic boundary dysfluencies and a segmentation method that labeled each word as essential or non-essential according to the minimum semantic and grammatical utterance extracted from the written sentences (p -value < $2.2e-16$). In general, substitute words in word substitutions were a) phonetically/graphically similar to the original words and b) more frequent in the Brazilian Portuguese language than original words. Self-corrections were relatively frequent (26%) in errors involving words (segmental errors) and virtually absent in errors involving prosody (suprasegmental errors, such as prosodic boundary insertions and pauses). Self-corrections of word substitutions and cacoeopies frequently involved the definite article or preposition preceding a substantive.*

Conclusions: *Oral reading errors are frequent among educated adults. The errors correlate with some linguistic patterns that may reflect fundamental sentence processing mechanisms. We develop hypotheses to explain our findings and offer some further speculations involving the concept of working memory. Additionally, we provide some tentative suggestions to improve written text elaboration in order to avoid oral reading errors.*

Keywords: *reading; fluency; prosodic boundaries; sentence processing; working memory*

LISTA DE FIGURAS

Figura 1— Aparência em tela do software de anotação de dados ELAN	76
Figura 2 – Fluxograma de análise do corpus	83
Figura 3 – Histograma de distribuição das idades	97
Figura 4 – Fluxograma da análise do corpus	98
Figura 5 – Percentual de períodos incongruentes por falante	104
Figura 6 – Número de períodos por extensão (em palavras)	105
Figura 7 – Extensão (em número de palavras) dos períodos C e I	106
Figura 8 – Média (por falante) do número de palavras por período congruente (C)	110
Figura 9 – Média (por falante) do número de palavras por período incongruente (I)	110
Figura 10 – Número médio de incongruências vs. tamanho do período: plotagem e linha de melhor ajuste	112
Figura 11 – Número médio de incongruências segmentares vs. tamanho do período: plotagem e linha de melhor ajuste	113
Figura 12– Número médio de incongruências suprasegmentares vs. tamanho do período: plotagem e linha de melhor ajuste	114
Figura 13 – Número de períodos por extensão (em palavras)	115
Figura 14 – Distribuição das taxas de elocução (palavras/minuto)	117
Figura 15 – Correlação entre taxa de elocução (palavras por minuto) e percentual de incongruências (número de períodos incongruentes/número total de períodos lidos)	120
Figura 16 – Idade vs. Percentual de incongruências	123

Figura 17 – Tempos entre respirações	126
Figura 18 – Número de palavras entre respirações	127
Figura 19 – Correlação entre “taxa de elocução” e “número de palavras entre respirações”	128
Figura 20 – Correlação entre “taxa de elocução” e “tempo entre respirações”	129
Figura 21 - Correlação entre tempo entre respirações e palavras entre respirações	129
Figura 22 – Número de ocorrências por tamanho da palavra incongruente	145
Figura 23 – Número de ocorrências por tamanho da palavra imediatamente anterior à palavra incongruente.....	146
Figura 24 – Histograma de frequências das palavras originais e inseridas.....	150
Figura 25 – Palavras originais e inseridas mais frequentes e o número de vezes em que aparecem nos textos e nas falas lidas.....	154
Figura 26 – Histogramas mostrando a distribuição das palavras por categoria.....	155
Figura 27 – Palavras aleatórias vs. inseridas: distribuição em categorias	159
Figura 28 – Semelhanças entre palavras: originais vs. inseridas e originais vs. aleatórias	160
Figura 29 – Incongruências suprasegmentares (ordem crescente de frequência)	169
Figura 30 – Percentual de cada posição hierárquica em períodos congruentes e incongruentes.....	180
Figura 31 - Modelo dual-route approach de processamento de leitura.....	215
Figura 32 - Embedded- processes approach to WM.....	248
Figura 33 – Aglutinação de palavras em unidades semânticas.....	264
Figura 34 – Um evento complexo, com múltiplos agentes e ações simultâneas	265
Figura 35 – Hipótese de possível período livre das limitações da WM.....	266

LISTA DE TABELAS

Tabela 1 - Unidades prosódicas vs. discursivas vs. sintáticas	89
Tabela 2 - Características demográficas dos falantes no momento da fala	96
Tabela 3 - Duração das falas (horas, minutos e segundos - hh:mm:ss)	99
Tabela 4 - Número de períodos com e sem incongruências	101
Tabela 5 - Percentual de períodos com e sem incongruências por falante, ordenados pelo percentual de incongruências	103
Tabela 6 - Número de palavras e de períodos por tipo de período	107
Tabela 7 - Palavras por tipo de período	109
Tabela 8 - Taxa de elocução (v) em palavras/minuto, em ordem crescente	116
Tabela 9 - Percentual de incongruências vs. taxa de elocução, ordenado segundo a taxa de elocução	119
Tabela 10 - Idade versus percentual de incongruências, ordenado pelo percentual de incongruências	122
Tabela 11 - Padrões respiratórios por falante	125
Tabela 12 - Todas as incongruências do corpus	131
Tabela 13 - Todas as incongruências do corpus, ordenadas por classificação	132
Tabela 14 - Incongruências segmentares	133
Tabela 15 - Categorias gramaticais nas incongruências de concordância	133
Tabela 16 - Desvios de concordância	134
Tabela 17 - Incongruências segmentares com e sem prejuízo ao período	134

Tabela 18 – Retratações em incongruências segmentares com prejuízo ao período	136
Tabela 19 - Incongruências segmentares com prejuízo e não retratadas	137
Tabela 20 - Incongruências segmentares frasais vs. lexicais	137
Tabela 21 - Retratações com ou sem adição de palavras não incongruentes.....	138
Tabela 22 - Categoria gramatical das palavras incongruentes retratadas, com e sem adição de palavra(s) congruentes(s) à retratação	139
Tabela 23 - Categoria gramatical da palavra retratada com ou sem adição de palavra(s) congruentes(s) à retratação	140
Tabela 24 - Número de palavras adicionadas à retratação, quando houve adição de palavra(s) congruente(s) à palavra incongruente	141
Tabela 25 - Categoria gramatical da palavra imediatamente anterior à palavra incongruente retratada, incluindo as situações em que a palavra foi adicionada e não foi adicionada à retratação.....	141
Tabela 26 - Categorias gramaticais das palavras incongruentes e das palavras adicionais incluídas nas retratações	142
Tabela 27 - Distribuição do número de palavras por tamanho (em sílabas) nas situações de retratação "com" e "sem" adição de palavras adicionais	143
Tabela 28 - Tamanho dos substantivos incongruentes nas condições "com" e "sem" adição de palavras à retratação	147
Tabela 29 - Tamanho das preposições (com ou sem art. def.) anteriores às palavras incongruentes nas condições em que foram ou não foram adicionadas à retratação.....	147
Tabela 30 - Trocas de palavras e retratações por falante	149
Tabela 31 – Número de trocas vs. Ocorrências nos textos vs. Ocorrências no CB	152
Tabela 32 - Categorias das palavras originais e inseridas.....	156

Tabela 33 - Categorias das palavras originais e inseridas.....	156
Tabela 34 - Trocas de palavra com troca de categoria.....	157
Tabela 35 - Palavras originais e inseridas com troca de categoria.....	157
Tabela 36 - Semelhança entre palavra original e inserida por tipo de palavra	158
Tabela 37 - Relação entre palavra original e aleatória* por tipo de palavra.....	160
Tabela 38 - Grau de concordância entre pesquisador e voluntários	166
Tabela 39 - Resultados dos graus de concordância entre pesquisador e voluntários.....	166
Tabela 40 - Concordância entre o pesquisador e o voluntário 1 na marcação do local de fronteiras inadequadas.....	166
Tabela 41 - Incongruências suprasegmentares por falante.....	168
Tabela 42 - Incongruências vs. <i>discourse segments</i>	170
Tabela 43 – Incongruências de fronteiras prosódicas vs. continuidade temática	171
Tabela 44 - Subtipos de inserção/deleção de fronteiras vs. <i>discourse segments</i>	171
Tabela 45 - Classificação dos atores (AIS) em períodos congruentes (C) e com incongruências (I) do tipo inserção/deleção de fronteira	173
Tabela 46 - Dados da Tabela 45 ordenados pela razão entre Períodos I e Períodos C.....	174
Tabela 47 - Sujeitos desinenciais ou de voz ativa ou passiva em períodos congruentes (C) e com incongruências (I) do tipo inserção/deleção de fronteira.....	175
Tabela 48 - Classificação dos sujeitos em períodos C e I.....	175
Tabela 49 - Segregação/agrupamento dos dados da Tabela 28	176
Tabela 50 - Pessoa/Posição vs. Congruência	177

Tabela 51 – Classificação hierárquica das palavras em períodos congruentes e incongruentes (inserção/deleção de fronteiras)	179
Tabela 52 – Classificação hierárquica das palavras em períodos congruentes e incongruentes (inserção/deleção de fronteiras)	180
Tabela 53 – Classificação hierárquica das palavras em períodos congruentes e incongruentes: ocorrências observadas/ocorrências esperadas	181
Tabela 54 - Relação hierárquica entre palavras adjacentes em períodos incongruentes	183
Tabela 55 - Relação hierárquica entre palavras: percentuais de ocorrência de incongruência prosódica	185
Tabela 56 - Tipo de relação hierárquica vs. inserção/deleção de fronteiras	186
Tabela 57 - Tipo de relação hierárquica vs. congruência e incongruência de períodos	187
Tabela 58 - Relações hierárquicas com maior incidência relativa de incongruências de inserção/deleção de fronteira	188
Tabela 59 - Relações hierárquicas: "desnivelamento" vs. congruência.....	189
Tabela 60 - Incongruência: posição no período	189
Tabela 61 - Incongruências: comparando as posições no período.....	190
Tabela 62 - Incongruências: posição relativa à predicação primária	191
Tabela 63 - Número de fixações por palavra	243

SUMÁRIO

APRESENTAÇÃO	18
1. INTRODUÇÃO	19
2. REVISÃO DE LITERATURA	24
2.1 A linguística como ciência	24
2.2 Teorias/modelos de processamento da linguagem: como escolher o(s) seu(s)?	26
2.3 O processamento de sentenças	28
2.3.1 O que processa?	29
2.3.2 Como processa?	30
2.3.3 Para que processa (qual o objetivo do processamento)?	33
2.4 Processamento de sentenças: alguns modelos emblemáticos.....	35
2.5 Evidências do cérebro em funcionamento: eletrofisiologia e neuroimagem funcional..	38
2.5.1 Das evidências neurobiológicas aos modelos cognitivos (e vice-versa): linhas tortuosas	40
2.5.2 Modelos cognitivos: alguns problemas fundamentais.....	40
2.6 A(s) memória(s) — limitações e consequências linguísticas	42
2.6.1 O papel da antecipação no processamento de sentenças	45
2.7 <i>Chunking</i> , <i>entrenchment</i> e outros modelos cognitivos.....	47

2.8 Constituintes sintáticos, estrutura argumental, argumentos, papéis temáticos, relações gramaticais e outras abstrações correlatas	51
2.9 A Prosódia (e outros fenômenos suprasegmentares)	60
2.10 Retomando alguns conceitos — prosódia, constituintes sintáticos, estrutura argumental, <i>chunking</i> — e vendo se é possível empacotá-los em outro: o <i>closure positive shift</i>	66
2.11 O corpus e as incongruências: fala lida, fala espontânea, fala lida espontânea.....	68
2.11.1 Especificidades do processo de leitura e suas consequências para a fala lida.....	69
3. METODOLOGIA	71
3.1 O corpus.....	73
3.2 Os falantes	73
3.3 Os critérios de seleção dos pronunciamentos analisados	74
3.4 Características do corpus e dos falantes após aplicação dos critérios de seleção.....	74
3.5 A identificação das incongruências	75
3.6 As incongruências.....	77
3.7 A validação das incongruências identificadas pelo pesquisador	84
3.8 Descrição do ambiente linguístico onde ocorreram (e não ocorreram) incongruências.	84
3.8.1 Segmentação do trecho do discurso em que se insere a sentença que contém a incongruência, inspirado no esquema de segmentação de estruturas linguísticas em segmentos de discurso (<i>discourse segments</i> — DS) proposto por Grosz e Sidner (1986, p. 183).....	85
3.8.2 Determinação do grau de proeminência do ator do período, segundo o modelo AIS (<i>actor identification strategy</i>) proposto por Bornkessel-Schlesewsky e Schlewski (2014).	85

3.8.3 Medição da extensão (em palavras) das sentenças onde ocorreram as incongruências.	86
3.8.4 Determinação da taxa de elocução (em palavras por minuto) no trecho em que ocorreu a incongruência.	87
3.8.5 Análises das trocas de palavras	87
3.8.6 Classificação hierárquica das palavras com relação à “predicação primária” (em períodos com incongruência por inserção/deleção inadequada de fronteiras prosódicas).	89
3.9 Aspectos complementares e coda metodológica	94
4. RESULTADOS	95
4.1 Falantes, corpus e incongruências <i>lato sensu</i>	95
4.1.1 Os falantes	95
4.1.2 O corpus	97
4.1.3 As incongruências	98
4.1.3.1 Períodos congruentes e incongruentes	98
4.1.3.2 Congruência vs. extensão dos períodos	104
4.1.3.3 Congruência vs. taxa de elocução	115
4.1.3.4 Congruência vs. idade	121
4.1.3.5 Congruência vs. sexo	123
4.1.4 Os padrões respiratórios	123
4.1.5 Resumo da seção 4.1	130
4.1.6 Prólogo das próximas seções (4.2 e 4.3)	131

4.2	Descrições e análises das incongruências segmentares	132
4.2.1	As incongruências segmentares	133
4.2.2	As retratações	135
4.2.3	As trocas de palavra.....	148
4.2.3.1	A frequência das palavras nos corpora	148
4.2.3.2	As categorias gramaticais das palavras trocadas	154
4.2.3.3	As semelhanças entre palavra original e inserida.....	158
4.2.4	Resumo da seção 4.2 (incongruências segmentares).....	161
4.3	As incongruências suprasegmentares.....	162
4.3.1	Descrição das incongruências suprasegmentares.....	162
4.3.2	Validação das incongruências do tipo inserção/deleção de fronteiras	164
4.3.3	Descrição geral das incongruências suprasegmentares identificadas	167
4.3.4	Incongruências suprasegmentares vs. “variáveis independentes”	169
4.3.4.1	Incongruências suprasegmentares vs. segmentos do discurso	169
4.3.4.2	Incongruências suprasegmentares vs. “estratégia de identificação do autor”	171
4.3.5	Resumo da seção 4.3	177
4.4	Predicação primária: palavras/constituintes obrigatórios vs. opcionais	178
4.4.1	Resumo dos resultados desta seção:	191
5.	DISCUSSÃO	192

5.1 Os falantes, o corpus e as incongruências <i>lato sensu</i>	193
5.1.1 Os falantes	193
5.1.2 O corpus.....	194
5.1.3 As incongruências.....	195
5.1.3.1 Os períodos congruentes e incongruentes	195
5.1.3.2 Congruência vs. idade.....	195
5.1.3.3 Congruência vs. sexo	196
5.1.3.4 Congruência vs. extensão dos períodos (e comentários adicionais).....	197
5.1.3.5 Congruência vs. taxas de elocução	199
5.1.3.6 Os padrões respiratórios	201
5.1.3.7 Incongruências segmentares vs. suprasegmentares	203
5.2 As incongruências segmentares	206
5.2.1 As retratações	208
5.2.2 As trocas de palavras	211
5.2.2.1 A frequência das palavras nos corpora	211
5.2.2.2 As semelhanças entre palavras originais e inseridas	214
5.2.2.3 Observações adicionais sobre as trocas de palavras	219
5.2.2.3.1 “Resistência” à substituição.....	219
5.2.2.3.2 O que é que o que é?.....	219
5.2.2.3.3 Tendência para ser escolhida como palavra substituta	223

5.2.2.3.4 O papel das categorias gramaticais nas trocas de palavras.....	223
5.3 As incongruências suprasegmentares.....	224
5.3.1 As incongruências de fronteiras prosódicas	225
5.3.1.1 Por que as incongruências de fronteira não são objeto de mais estudos?.....	226
5.3.1.2 Como foram validadas	226
5.3.1.3 Por que as incongruências suprasegmentares não são retratadas?	227
5.3.1.4 Incongruências de fronteira prosódica: correlação com a segmentação do discurso	229
5.3.1.5 Incongruências de fronteira prosódica: correlação com a estratégia de identificação do ator.....	230
5.4 Predicação primária e hierarquia argumental	232
5.5 Discussões adicionais	240
5.5.1 A natureza das fronteiras prosódicas	240
5.5.2 Fala espontânea vs. “fala escrita”	241
5.5.3 Produção & compreensão	245
5.5.4 A incontornável memória de trabalho (WM)	246
5.5.5 Hipóteses adicionais	260
5.5.5.1 Hipótese I.....	260
5.5.5.2 Hipótese II	262
5.5.5.3 Hipótese III	266
5.6 Sugestões para a evitar incongruências de fala lida.....	269

5.6.1 Incongruências segmentares	271
5.6.2 Incongruências suprasegmentares	271
5.6.3 Sugestão adicional (e final): sobre a escrita de números	277
6. CONCLUSÃO	278
7. REFERÊNCIAS	280

APRESENTAÇÃO

O título deste trabalho — *Incongruências em fala lida: estudo de corpus de pronunciamentos parlamentares* — sofre de certo comedimento, beirando o eufemismo. Infelizmente, por questões burocráticas relacionadas aos passos necessários à defesa da dissertação, o título não pode ser modificado no preciso momento em que o excesso de cautela se revelou como falta de clareza. Sendo assim, fez-se necessário acrescentar essa breve “apresentação” para esclarecer melhor o que o título atual, fixado a contragosto do autor, não esclarece. O título desejado teria sido algo como “erros de fala lida: estudo de corpus em adultos escolarizados”; ou, talvez, para atrair mais leitores, “dificuldades na fala lida: quem são, onde vivem e do que se alimentam?”. No entanto, o título original permanece, e o esclarecimento abaixo busca suprir essa falha.

O trabalho estudou um corpus composto por vídeos de pronunciamentos parlamentares — discursos políticos — proferidos por indivíduos adultos e escolarizados, identificando e analisando incongruências entre as leituras em voz alta e os respectivos textos escritos em que elas se basearam.

Incongruências de todos os tipos — incluindo disfluências típicas, fronteiras prosódicas, e demais divergências entre texto escrito e leitura — foram analisadas e correlações estatísticas com variáveis linguísticas e demográficas foram estabelecidas.

Com base nos resultados do estudo, levantaram-se algumas hipóteses para tentar compreender e explicar os achados.

1. INTRODUÇÃO

Em um de seus diálogos, Platão ([ca 360 a.C.]) descreve uma conversa entre dois personagens: Fedro e Sócrates. Fedro acabara de ouvir um discurso sobre o “amor”, e Sócrates, argumentando que era apaixonado pela retórica, pede ao companheiro que lhe repita o que ouvira. Fedro resiste, diz que não se lembra, mas Sócrates, suspeitando que o amigo trazia consigo uma cópia escrita do discurso, insiste até obter o que desejava.

Depois de ouvir a leitura, Sócrates conta a história de um deus egípcio chamado Teuto, que teria inventado a matemática, a aritmética, o cálculo, a geometria, a astronomia e — seu maior orgulho — a escrita. Teuto ofereceu suas criações a um determinado rei, que teria ficado bastante admirado e agradecido com tais maravilhas, mas teria implicado, justamente, com a escrita. O rei argumentou que a escrita era uma invenção muito ruim, mera reminiscência com enganosa aparência de verdade. E Sócrates, para arrematar seu raciocínio, acrescentou que, assim como a pintura — que pode parecer viva, mas está morta —, a escrita seria apenas uma imagem, um pálido reflexo da voz, algo que não fala, não responde, e nem é capaz de se defender.

Como tantas ideias gestadas naquela época — e que, ainda hoje, nos surpreendem por sua contemporaneidade —, essa interpretação da modalidade escrita permanece bastante atual. John Lyons (LYONS, 1982) diz, com propriedade, que fala e escrita são produtos não isomórficos e não equivalentes do comportamento linguístico. Oconnell e Kowal (2008) vão além, recuperando uma antiga metáfora segundo a qual a escrita seria apenas o cadáver da fala. Outros autores, entretanto, dão mais crédito à escrita e preferem estudar a possibilidade de ressuscitá-la, de aproximá-la da fala e trazê-la de volta à vida por meio da aplicação do conceito de fluência leitora. Entre esses, estão Daane et al. (2002), que definem fluência como "aderência à sintaxe do autor"; e Kuhn et al. (2010, p. 240), que a conceituam como "uma combinação de acurácia, automaticidade e (...) prosódia que, tomados em conjunto, facilitam a reconstrução do sentido pelo leitor; e que se demonstra, na leitura oral, por meio da facilidade de reconhecer as palavras e de imprimir ritmo, fraseamento e entoação adequados [ao texto]."¹ De fato, os conceitos de fluência vs. disfluência variam entre autores, havendo certa tradição na literatura fonoaudiológica, como se observa em Schiefer e Arcuri (2014), que limita bastante os

¹ “*Fluency combines accuracy, automaticity, and oral reading prosody, which, taken together, facilitate the reader’s construction of meaning. It is demonstrated during oral reading through ease of word recognition, appropriate pacing, phrasing, and intonation.*”

fenômenos analisados. Neste trabalho, empregamos conceitos mais abrangentes, inclusivos, considerando todas as manifestações em que o leitor se afasta da sintaxe ou da mensagem pretendida pelo texto escrito e denominando-as “incongruências” (entre texto escrito e fala lida).

Não é necessário recorrer a filósofos gregos ou a linguistas consagrados para constatar que leitura e fluência, ou congruência entre texto escrito e fala lida, de fato, não andam sempre juntas. Tampouco é necessário ir aos bancos escolares em busca de aprendizes com pouca experiência de leitura. A verdade é que até mesmo comunicadores tarimbados — como apresentadores de telejornais, cujas profissões dependem da leitura em voz alta — apresentam, ocasionalmente, dificuldades com a leitura fluente. Ou seja, leitores experientes, com dezenas de anos e, talvez, milhares de horas de leitura na bagagem, também estão sujeitos a disfluências de leitura. E esse fenômeno também pode ser observado — retomando, de certa forma, o diálogo entre Sócrates e Fedro — nos discursos políticos proferidos no Plenário do Senado Federal (BRASIL, Senado Multimídia, 2019). Não é infrequente que os Senadores se utilizem de textos escritos para fazer seus pronunciamentos. Tampouco é incomum que apresentem, durante a leitura de seus discursos, certas disfluências ou erros de leitura (incongruências entre texto escrito e fala lida) que tornam perceptível aos ouvintes que não se trata de fala espontânea, mas de falas lidas.

Naturalmente, não há qualquer objeção ao uso de textos escritos como meio auxiliar para proferir pronunciamentos. É recurso usual em discursos políticos em qualquer lugar do mundo, como comenta Lehrman (2010) em seu manual de elaboração para esse gênero textual. A principal razão para a necessidade de textos escritos é a limitação da memória humana para o armazenamento de falas que, muitas vezes, têm dezenas de minutos de duração. Independentemente do impasse teórico entre a possibilidade ou a impossibilidade de se registrar com precisão, por meio da escrita, uma mensagem que se pretenda ver transformada em fala, como defendeu Sócrates — ou, até mesmo, em virtude desse impasse —, pareceu-nos interessante observar — sob a ótica de conceitos mais amplos de fluência, como os elencados acima — as incongruências entre texto escrito e falado eventualmente presentes nesse banco de dados público disponibilizado pelo Senado Federal (BRASIL, Senado Multimídia, 2019).

E, considerando tal fenômeno — a dificuldade em trazer de volta à vida, à fala, sem sequelas, um texto escrito — e o fato de que a literatura não costuma estudar a congruência de leitura em indivíduos adultos — dando ênfase a crianças e jovens em idade escolar —, elaboramos as questões centrais que impulsionam este estudo: a) Qual será a prevalência de incongruências de leitura em falas

lidas produzidas por indivíduos adultos e escolarizados? b) Haverá correlação entre as incongruências observadas nesse corpus (BRASIL, Senado Multimídia, 2019) e aspectos linguísticos específicos dos textos escritos? Em outras palavras: considerando “incongruências” de leitura como leituras que não correspondem ao texto escrito e que, de alguma forma, prejudicam a comunicação das ideias contidas no texto, será que a análise de aspectos estruturais (morfofossintáticos, semânticos e discursivos) dos textos escritos revelará características linguísticas específicas que impliquem uma maior incidência de incongruências na leitura desses textos?

Acreditamos que a resposta a essas perguntas deve ser afirmativa e, portanto, nossas **hipóteses iniciais** são que:

- a) determinados aspectos morfofossintáticos, semânticos e/ou discursivos dos textos escritos devem estar associados a maior ou menor frequência de incongruências de leitura
- b) tais correlações devem refletir aspectos fundamentais dos mecanismos de processamento de linguagem

Para testar essas hipóteses, estabelecemos como **objetivos gerais**:

- a) correlacionar incongruências de leitura com aspectos estruturais do texto escrito
- b) identificar, na análise dessas correlações, possíveis mecanismos fundamentais de processamento da linguagem

E, para atingir tais objetivos gerais, elaboramos uma série de **objetivos específicos**, a saber:

- a) identificar, sistematicamente, as incongruências mais comuns em nosso corpus de falas lidas (pronunciamentos parlamentares)
- b) identificar, nos textos escritos que deram origem às falas lidas, determinadas características linguísticas (morfofossintáticas, semânticas e discursivas) pré-selecionadas
- c) analisar as correlações estatísticas eventualmente existentes entre os itens “a” (incongruências) e “b” (características linguísticas)
- d) analisar os achados de “c” (correlações estatísticas) à luz de teorias atuais de processamento da linguagem

- e) discutir se os achados analisados em “d” poderiam, eventualmente, contribuir para o aperfeiçoamento de teorias atuais de processamento de sentenças, ou de teorias cognitivas de um modo geral
- f) elaborar, com base nos achados do estudo, possíveis recomendações para o aperfeiçoamento do processo de produção de textos escritos cujo objetivo seja a leitura e voz alta, especialmente quando relacionada ao fonostilo político.¹

O texto a seguir organiza-se nos capítulos habituais — revisão de literatura, metodologia, resultados, discussão, conclusão e referências. O próximo capítulo, a revisão de literatura, fatiará o corpo teórico relacionado ao nosso problema de pesquisa e o organizará sob os seguintes tópicos:

2.1 Algumas questões relacionadas à linguística como ciência

2.2 Defesa de uma abordagem neurobiológica da linguagem

2.3 O que se estuda no campo do processamento de sentenças: quem processa o quê, como e para quê?

2.4 Alguns modelos emblemáticos de processamento de sentenças

2.5 As principais ferramentas neurofisiológicas utilizadas nas investigações de processamento de sentenças, e algumas considerações sobre como os diversos modelos cognitivos interpretam de forma diversa os achados obtidos com o uso dessas ferramentas

2.6 A importância da memória nos processos linguísticos e alguns modelos cognitivos que consideram suas limitações no processamento de sentenças

2.7 Alguns mecanismos propostos para contornar dificuldades de processamento

2.8 Estrutura argumental, constituintes sintáticos e outras hipóteses de organização estrutural das sentenças

¹ Como bem apontam Lima e Constantini (2017), o fonostilo político caracteriza-se por possuir caráter mais persuasivo; entoação com frequência fundamental com maior variação entre frequências altas; e ajustes desviantes e mudanças nas características da qualidade vocal associados às atitudes e às emoções que se pretende expressar.

2.9 O papel da prosódia no processamento de sentenças

2.10 O *closure positive shift* como uma manifestação eletrofisiológica cerebral que pode estar relacionada a certas hipóteses de segmentação de sentenças e de segmentação cognitiva em geral

2.11 Especificidades do processo de leitura (objeto de nosso corpus) em relação à fala espontânea

Os demais capítulos — metodologia, resultados, discussão, conclusão e referências — seguem o padrão habitual e seus conteúdos correspondem a seus nomes de batismo.

2. REVISÃO DE LITERATURA

2.1 A linguística como ciência

No estudo da linguagem, que grau de objetividade é possível atingir, e que grau de subjetividade é possível e desejável evitar?

Lyons (1981), no Capítulo 2 de seu clássico *Language and linguistics: an introduction*, lança a seguinte pergunta: a linguística é uma ciência? Não se trata de uma pergunta retórica, pois o autor desenvolve sua resposta com suficientes detalhes. Apesar de uma revisão pormenorizada dessa questão não ser condizente com o escopo deste estudo, parece útil mencionar, de passagem, as formas pelas quais a linguagem e as línguas têm sido, ou podem ser, cientificamente abordadas.

Dos conceitos tratados por Lyons (1981), interessam-nos o empirismo, o reducionismo e o racionalismo. O ponto de partida da discussão do autor é a própria linguagem, ou a maneira como a linguística é capaz de estudá-la. Segundo ele, sob o ponto de vista dos estudos científicos, a linguagem e as línguas, transformadas em objetos de estudo, seriam melhor interpretadas como construtos teóricos, modelos abstratos cujo objetivo seria descrever e explicar, da melhor maneira possível, os fenômenos linguísticos observados.

O empirismo, o primeiro conceito, privilegiaria a observação direta dos fenômenos naturais por meio dos sentidos mais básicos, sem a interferência de hipóteses ou de interpretações pré-concebidas. O comportamento linguístico seria um desses fenômenos naturais, e hipóteses e explicações linguísticas deveriam ser formuladas por meio do raciocínio indutivo, a partir das observações primárias.

O conceito seguinte, o reducionismo — discutido por Lyons (1981) sob o ponto de vista do positivismo do Círculo de Vienna, em sua vertente fisicalista (THOMAS UEBEL, 2017) —, defendia que, em última instância, todos os fenômenos observados deveriam ser reduzidos a seus aspectos físicos palpáveis. Um fenômeno cultural, por exemplo, deveria ser reduzido, progressivamente, a seus aspectos sociológicos, psicológicos, biológicos, físicos e matemáticos para que pudesse ser considerado digno de uma verdadeira teoria científica. Lyons diz que essa abordagem, assim como o empirismo radical, vem sendo abandonada, pois cada disciplina teria seus próprios métodos de estudo e, no caso da linguística, uma abordagem reducionista seria "estéril" e "inviável".

Lyons prossegue citando Karl Popper e demonstrando as vantagens de um empirismo atenuado, pois seria improvável e contraproducente empreender uma jornada em busca de dados sem ter, previamente, formulado algumas hipóteses. Nessa linha de raciocínio, lembramos que Popper (THORNTON, 2018) propunha que as descobertas científicas se originariam dos problemas com os quais nos deparamos, e que a busca por soluções para esses problemas é que nos levaria à observação da natureza e à formulação de hipóteses e teorias. O conhecimento científico, portanto, não seria derivado, majoritariamente, do raciocínio indutivo, mas do raciocínio dedutivo. Popper retomou, então, a importância da mente como um laboratório de experimentos e de ideias, e propôs um racionalismo crítico, em oposição ao empirismo puro. De acordo com sua tese, o conhecimento evoluiria de teoria em teoria, todas elas provisórias, e as teorias mais atuais estariam sempre expostas ao falseamento definitivo e/ou à substituição por teorias alternativas, mais adequadas à explicação dos fenômenos observados.

No campo da linguística, o maior expoente da abordagem racionalista — proponente da chamada "concepção racionalista do fenômeno da linguagem" — talvez seja Noam Chomsky (UEBEL, 2017). Chomsky defende o uso da intuição linguística — de nossa competência como falantes de línguas particulares — como instrumento genuíno para compreender fenômenos linguísticos universais. E isso nos leva de volta a Lyons (1981), que aborda a intuição do falante e, em especial, a intuição do linguista, como um passo inevitável e não necessariamente indesejável no caminho dos estudos da linguagem. A introspecção do linguista, portanto, quando exercida sob a ótica do racionalismo crítico de Popper, seria parte essencial do instrumental científico à disposição da ciência linguística.

Para finalizar essa breve incursão por aspectos da filosofia da ciência — cuja justificativa será dada a seguir —, mencionaremos a visão que Lakoff e Johnson (2003) descreveram em sua obra *Metaphors we live by*. Eles modulam nossas pretensões objetivistas e defendem que a grande questão da abordagem científica é, simplesmente, evitar o viés individual (*individual biases*, no original). Segundo eles, tanto o objetivismo quanto o subjetivismo seriam mitos — num sentido não pejorativo — criados pela filosofia ocidental. Não parece necessário dizer que o subjetivismo puro deve ser evitado; mas seria interessante citar a visão dos autores sobre o objetivismo. Segundo eles, a verdade absoluta seria inatingível, pois seria sempre interpretada a partir do sistema conceitual e metafórico que alicerça nossa cognição. E esse sistema conceitual e essencialmente metafórico, por sua vez, estaria enraizado em — e seria intimamente dependente de — nosso corpo e nosso sistema nervoso.

Essa interpretação — chamada *embodied cognition* — dá origem à teoria neural do pensamento e da linguagem (*neural theory of thought and language* — NTTL), que Lakoff (2012) resume da seguinte forma: a) nós pensamos por meio de nossos cérebros, o que significa que o pensamento é físico e realizado por circuitos neuronais; b) o que faz o pensamento ter sentido é a forma como os circuitos neuronais se conectam ao corpo; e c) tanto as ideias abstratas quanto a linguagem são, assim, corporificadas. Em suma: a ideia de que nossa cognição manipula objetivamente símbolos abstratos independentes, desconectados da realidade corporal e neural (onde a manipulação acontece), não faria o menor sentido. A realidade externa, segundo essa visão, seria percebida por intermédio de sua representação interna (corporal e cerebral) e seria irremediavelmente dependente dela.

2.2 Teorias/modelos de processamento da linguagem: como escolher o(s) seu(s)?

Nosso ponto de partida: abordar a linguagem como um fenômeno cognitivo de base biológica.

A justificativa para termos empreendido essa arriscada aventura por diferentes abordagens sobre a ciência em geral e a linguística em particular é que as teorias e os modelos que tentam compreender o processamento da linguagem — a forma como processamos e compreendemos nossa língua, tanto falada, quanto escrita — são, também, bastante variados em seus pressupostos teóricos e suas conclusões.

A divergência nas abordagens teóricas do fenômeno do processamento da linguagem lembra a fábula dos cegos que tentam descrever um elefante: ou algum dos teóricos não é cego (e atingiu uma visão realmente próxima da realidade do fenômeno linguístico), ou vários deles podem ter, eventualmente, capturado grãos de realidade em seus pressupostos e/ou conclusões. E como este estudo visa, como descrito em sua introdução, a “correlacionar incongruências de leitura com aspectos estruturais do texto escrito e identificar, na análise dessas correlações, possíveis mecanismos fundamentais de processamento da linguagem”, será necessário escolher pontos de partida teóricos para essas análises. Teremos, portanto, de selecionar, a priori, teorias e modelos de processamento da linguagem que nos pareçam mais convincentes, convenientes e produtivos.

Um parêntesis torna-se necessário neste momento para estabelecer uma diferença entre os conceitos de "modelo" e "teoria". Poderíamos considerar teoria como uma explicação concisa, coerente, sistemática, preditiva e amplamente aplicável a um conjunto de fenômenos delimitados; e modelo, como uma abordagem com viés computacional e matemático dos fenômenos explicados pela

teoria (THE UNIVERSITY OF CALIFORNIA MUSEUM OF PALEONTOLOGY, 2019). Mas, apesar de haver tentativas de estabelecer diferenças precisas entre esses conceitos, diferentes disciplinas, com seus diferentes métodos de pesquisa, usam diferentes critérios (GOLDFARB; RATNER, 2008). Sendo assim, optaremos por empregar ambos — teoria e modelo — de forma intercambiável, usando, sempre que possível, a terminologia escolhida pelo autor da teoria ou do modelo.

Um dos critérios que aplicaremos, então, na seleção do modelo/teoria será o reducionismo. Não o reducionismo fisicalista de inspiração vienense, mas um reducionismo que parte do pressuposto de que a linguagem é um fenômeno **cognitivo** de base **biológica**, cuja existência depende de um substrato neurobiológico. A diferença entre "perspectiva cognitiva" — ponto de partida dos modelos abstratos de processamento da sintaxe, da semântica e do léxico — e "perspectiva neurobiológica" — como nos permitem certas ferramentas investigativas como potenciais evocados, magnetoencefalografia e exames de imagem funcionais, por exemplo — seguem a abordagem proposta por Bornkessel-Schlesewsky, Staub e Schlewsky (2016). Assumiremos, então, sempre que possível, uma perspectiva de **coerência neurobiológica** na seleção dos modelos cognitivos de processamento da linguagem. Ou seja, toda e qualquer teoria/modelo considerado deverá ser capaz de prestar contas ao substrato neurobiológico onde o fenômeno linguístico acontece.

E, como nossas observações empíricas iniciais (que motivaram este estudo) revelaram que as incongruências entre a escrita e a fala mais comumente identificadas podem ser divididas em dois tipos — segmentares (palavras) e suprasegmentares (prosódia) —, selecionaremos modelos capazes de explicar tais incoerências. Apesar de concordarmos com Sinclair (2001, p. 357, apud IZREEL, 2018), quando diz que *"to me, a corpus of any size signals a flashing neon sign 'think again', and I find it extremely difficult to fit corpus evidence into received receptacles... the language obstinately refuses to divide itself into categories prepared in advance for it"*¹; e com Blanche-Benveniste e JeanJean (1987, p. 90, apud IZREEL, 2018), que afirmam que *"it is not about using [language X] to illustrate a theory, but finding a theory that allows to approach the data of spoken [language X]"*²; apesar disso,

¹ Para mim, qualquer corpus acende uma luz de alerta dizendo “pense novamente”, e acho extremamente difícil acomodar dados extraídos de corpora em receptáculos pré-estabelecidos... a língua recusa-se a se encaixar em categorias previamente estabelecidas para ela. (tradução nossa)

² Não se trata de usar a [língua X] para ilustrar uma teoria, mas de encontrar uma teoria que nos permita entender os dados da [língua X] (tradução nossa)

não nos resta alternativa senão partir de algum modelo específico, ou da adaptação de certos modelos coerentes entre si, com seus respectivos receptáculos e ilustrações, para analisar e explicar os dados linguísticos observados.

A seguir, discutimos brevemente quais elementos fazem parte do jogo no campeonato de modelos/teorias; e, na sequência, abordamos alguns modelos específicos.

2.3 O processamento de sentenças

'Estrutura argumental' do verbo PROCESSAR (sentenças): quem processa o quê (2.3.1), como (2.3.2), e para quê (2.3.3)?

Os diversos modelos de processamento de sentenças pretendem descrever, partindo de pressupostos e evidências empíricas e teóricas, como os indivíduos processam e compreendem uma sentença (ou um enunciado, ou um discurso, ou um diálogo, ou um texto escrito, ou, mesmo, em modelos mais recentes, as línguas de sinais). O ponto de partida para a visão geral dos modelos apresentada nesta seção é uma revisão recente feita por Ferreira e Çokal (2016). A partir dessa análise esquemática, acrescentaremos pontos de vista de autores adicionais que, segundo nossa análise, parecem trazer contribuições interessantes.

O primeiro aspecto a ser considerado é o próprio indivíduo que se defronta com uma sentença e pretende compreendê-la. Na literatura inglesa, o nome "*parser*" tem sido usado, principalmente, para designar o "*human sentence comprehension mechanism*" (LOGACEV; VASISHTH, 2015). Ou seja, parte-se do pressuposto de que há um mecanismo cerebral dedicado a esse propósito. Em português brasileiro (PB), adota-se também, com muita frequência, a mesma terminologia (RIBEIRO, 2015; MAIA; FINGER, 2005). Nivre (2005), ajuda-nos a compreender melhor a origem desse termo e a tomar certas decisões com relação a seu uso. Segundo esse autor, *parsing* deriva do latim *pars*, de *pars orationis*, que deu origem ao inglês *parts of speech*. Originalmente, *parsing* denominava a prática de analisar gramaticalmente as sentenças nas escolas primárias. Depois, foi adotado com conotação especializada pela linguística e pela ciência da computação com o sentido de "análise sintática de sentenças de um texto". O *parser*, portanto, no processamento de sentenças, seria, originalmente, o mecanismo que decomporia a sentença em suas partes, executaria uma análise sintática e, em seguida, uma ressíntese da sentença que deveria ser compreendida.

Como se verá a seguir, esse viés sintático do processamento de sentenças não atinge todos os modelos existentes, não justificando, senão por tradição, o uso do termo *parser*. Além disso, como não existe uma tradição em PB para o uso desse termo — apesar de seu uso na literatura dedicada ao tema, citada acima —, preferimos utilizar, neste trabalho, um termo de nossa própria língua, mais genérico e sem filiação teórica com modelos específicos. Usaremos, portanto, o termo "intérprete" em lugar de *parser*. Intérprete indicará, neste texto, não um mecanismo teórico específico dedicado ao processamento de sentenças, mas o indivíduo que enfrenta a tarefa de compreender uma determinada sentença de uma língua qualquer. Acreditamos que o emprego desse termo nos permitirá algum grau de liberdade em relação a qualquer pressuposto teórico prévio.

Portanto, a questão se resume a um intérprete que precisa processar certas informações linguísticas para compreender determinada sentença. Afastada essa questão inicial, a discussão da literatura está centrada, basicamente, em variações destes três elementos: a) qual ou quais informações (o que processa); b) são processadas de que forma (como processa); c) para atingir que grau de compreensão (para que processa).

2.3.1 O que processa?

Sintaxe, léxico, prosódia, contexto...

Considerando as abordagens teóricas que tomaram corpo a partir da década de 1980, podemos dizer que, inicialmente, havia uma atenção muito forte para os aspectos sintáticos. As informações que o intérprete, inicialmente, deveria obter seriam de caráter exclusivamente sintático e o objetivo inicial do processamento seria reconstruir fielmente a estrutura sintática original da sentença. A tarefa frente a uma sentença ambígua como "O rapaz viu a moça da moto" implicaria o processamento de regras básicas para a construção da árvore sintática adequada. Em uma segunda etapa, seriam interpretadas as informações semânticas que, se não fossem compatíveis com a estrutura sintática obtida, obrigariam o intérprete a reconstruir a árvore sintática da sentença. A eventual preferência do intérprete por uma ou outra construção sintática ("da moto" em função adverbial — "Da moto, o rapaz viu a moça", ou como adjunto de "a moça" — "O rapaz viu a moça da moto") deveria refletir regras sintáticas fundamentais, universais e imutáveis.

Ambiguidades como a que citamos poderiam induzir o intérprete a um erro interpretativo (sintático e semântico) inicial, um fenômeno que recebeu o nome de "*garden path*". Em PB, é comum o uso do termo "labirinto" em lugar de "*garden path*". Daí, veem-se expressões como "teoria do

labirinto", "modelo do labirinto" (RIBEIRO, 2015; MAIA; FINGER, 2005) e assim por diante. Essa terminologia leva a crer que a sentença é difícil de interpretar e que o intérprete ficaria perdido entre inúmeras pistas sintáticas capazes de levá-lo por caminhos diversos. A expressão em inglês, entretanto, segundo o *Oxford Dictionary of English Idioms* (AYTO, 2009), tem sentido diferente. Vem de "*lead someone down the garden path*", que significa seduzir, flertar, levar alguém por caminhos enganosos. Tal conotação parece mais adequada, pois as tais sentenças ambíguas indutoras do efeito "*garden path*", quando tal efeito é observado, parecem muito mais *enganosas* do que propriamente *labirínticas*. Neste texto, portanto, usaremos a expressão original — "*garden path*" — abreviada como GP. Daí, efeito GP, teoria GP etc., em vez de efeito ou teoria do labirinto.

Em oposição às teorias que consideravam que o intérprete processava, inicialmente, apenas as informações sintáticas, surgiram teorias lexicalistas, que punham as palavras, em vez da sintaxe pura, em lugar de destaque. A estrutura sintática não seria construída a partir de regras formais aplicáveis às categorias gramaticais das palavras, mas a partir das próprias palavras, que já conteriam, em si mesmas, as regras aplicáveis a cada uma delas. E, como extensão lógica dessa abordagem, informações semânticas, como número e animacidade, passaram a ser consideradas por teorias lexicalistas na interpretação inicial da sentença.

Nesse mesmo período, além das teorias sintáticas e lexicais, começaram a ganhar força certas correntes que defendiam a importância de informações advindas de outros aspectos linguísticos, como a prosódia e o contexto discursivo, no processamento inicial de sentenças. O panorama que se observa, portanto, compõe-se de diversos estudos buscando evidências empíricas da influência de cada uma dessas fontes de informação — sintaxe, léxico, prosódia e contexto — no processamento de sentenças, e gestando teorias e modelos para explicar os dados, não necessariamente nessa ordem. Os resultados desses estudos são variados, às vezes complementares, mas muitas vezes conflitantes.

2.3.2 Como processa?

Processamento simultâneo, atrasado ou adiantado? Serial ou paralelo? Modular ou interativo?

Lembrando que a discussão da literatura está centrada em variações destes três elementos — a) qual ou quais informações; b) são processadas de que forma; c) para atingir que grau de compreensão —, passaremos ao item “b”, seguindo a organização temática da revisão de Ferreira e Çokal (2016)

Se há controvérsias na definição da importância de cada fonte de informação no processamento de sentenças, a forma como essas informações são processadas gera controvérsias ainda mais profundas. Os diversos modelos divergem a respeito de questões que podem, grosso modo, ser divididas em quatro categorias:

- descompassos na interpretação incremental
- processamento serial ou paralelo
- modularidade ou interatividade
- fontes de complexidade no processamento.

A “interpretação incremental” sugere que o intérprete constrói o sentido da sentença ao mesmo tempo em que vai recebendo as informações, palavra por palavra. Há modelos que apregoam um descompasso nessa interpretação, tanto antecipando (predizendo) informações ainda indisponíveis, quanto retardando a interpretação em determinadas situações. Algum grau de predição já foi demonstrado para palavras lexicais (ALTMANN; KAMIDE, 1999, apud FERREIRA; ÇOCAL, 2016) e gramaticais (STAUB; CLIFTON, 2006, apud FERREIRA; ÇOCAL, 2016). E certo *delay*, um retardo no prosseguimento da interpretação, também foi demonstrado sob a forma de uma lentificação de leitura ao final de frases e sentenças, ou de um retardo na decisão do intérprete entre palavras com mais de um sentido possível ou construções sintáticas concorrentes (o rapaz viu a moça da moto). Nesses casos, o intérprete retarda seu comprometimento com este ou aquele sentido até obter informações suficientes para a desambiguação.

O segundo foco de divergência é o modo de processamento: serial ou paralelo. No processamento serial, as informações disponíveis são interpretadas uma após a outra, uma de cada vez, de forma seriada. No processamento paralelo, mais de uma informação pode ser processada ao mesmo tempo, em paralelo. As teorias que consideram o processamento serial costumam dar mais ênfase às informações sintáticas e à construção da estrutura sintática, passo a passo, seguindo regras específicas, como na sintaxe gerativa. As que consideram um processamento paralelo opõem-se, em princípio, ao paradigma gerativista.

O terceiro item — a oposição entre modularidade e interatividade — pode ser considerado, de certa forma, uma extensão do anterior. Os modelos modulares reportam-se à ideia de modularidade da linguagem, como na hipótese chomskiana de que há, no cérebro, um módulo específico dedicado à linguagem, ou, mais especificamente, à sintaxe. Segundo esses modelos, as informações linguísticas

(predominante ou exclusivamente sintáticas) são processadas com prioridade, no "modulo da linguagem", isoladas das demais informações disponíveis. Os que defendem a interatividade, por outro lado, acreditam que todas as informações — sintáticas, lexicais, discursivas, pragmáticas, contextuais ou quaisquer outras que estejam disponíveis ao intérprete naquele momento — são processadas ao mesmo tempo e se comunicam umas com as outras desde o início do processamento da sentença.

O quarto motivo de divergências diz respeito às diversas fontes de complexidade (de dificuldade) impostas ao processamento. Um subgrupo dessas complexidades são as chamadas ambiguidades, extensamente estudadas em condições experimentais controladas. De acordo com Trask (1996), ambiguidade é o fenômeno segundo o qual uma mesma palavra ou sequência de palavras podem evocar sentidos distintos. Ela pode ser lexical, como no caso dos homógrafos (“a forma do bolo era quadrada”, com dois significados possíveis: “o bolo é quadrado”, ou “o bolo foi feito numa vasilha quadrada”), homófonos (“vamos caçar/cassar os políticos”) ou homônimos (“o cão está mastigando a manga”: a manga da camisa ou a manga da mangueira?). Ou pode ser estrutural, como em:

(1) Ele parece o cão quando chupa manga.

(2) Suspeito que ele perseguia é capturado.

Em (1), a ambiguidade, além de lexical (manga vs. manga), é estrutural, pois a sentença tem duas interpretações possíveis: (a) Quando chupa manga, ele parece o cão; ou (b) Ele parece o cão quando [o cão] chupa manga. Além disso, é do tipo global, pois a ambiguidade é persistente, não se resolve com base, unicamente, na análise da estrutura da sentença. Em (2), a ambiguidade seria local (transitória), pois uma provável interpretação de "suspeito" como verbo de uma sentença *pro-drop* (~~eu~~ suspeito que ele perseguia...) pode ocorrer em determinado momento durante o processamento da sentença, mas é logo abandonada em benefício da interpretação correta de “suspeito” como substantivo.

Em (3), abaixo, temos outro caso de ambiguidade, tanto lexical, quanto estrutural global, conforme se interprete "ama" como verbo ou substantivo, acarretando a interpretação de "o guarda", respectivamente, como sintagma nominal ou verbal, e “velha” como substantivo ou adjetivo.

(3) A velha ama o guarda

Outra complexidade muito discutida é a das chamadas "*nested structures*", ou "*center-embedded sentences*", como a série de sentenças exemplificadas em Lobato (1986, p. 52):

(4) O rapaz saiu.

(5) O rapaz que o homem viu sai.

(6) O rapaz que o homem que a moça convidou viu saiu.

(7) O rapaz que o homem que a moça que João beijou convidou viu saiu.

Não há consenso quanto à fonte da crescente complexidade das sentenças de (4) a (7) — ou até mesmo da inaceitabilidade, apesar da absoluta normalidade gramatical, de sentenças como (6) e (7) — , mas postulam-se, entre outras possibilidades, limitações da memória de trabalho, ou a frequência relativa dessas estruturas na língua do dia a dia.

Um detalhe curioso é que sentenças como (7) podem ser reconstruídas com a mesma quantidade de informação por meio de uma estrutura ramificada à direita (no original, "*right-branching*"), sugerindo que o problema não seria a quantidade absoluta de informação, mas a forma como a informação foi estruturada sintaticamente:

(8) João beijou a moça que convidou o homem que viu o rapaz que saiu.

Ou, para manter a topicalidade de "rapaz":

(9) O rapaz que foi visto pelo homem que foi convidado pela moça que João beijou saiu.

2.3.3 Para que processa (qual o objetivo do processamento)?

Compreensão total e profunda, ou parcial e superficial?

Tendo mencionado brevemente quem, o que e como são processadas as informações linguísticas, passamos ao objetivo teleológico de todo o processo, o propósito final do intérprete, que seria, em tese, a compreensão da sentença. Os modelos iniciais partiam da premissa de que o intérprete envidava os melhores esforços para compreender integralmente a sentença, só ficando satisfeito quando saísse do processo — ou transferisse sua atenção para a próxima sentença ou atividade — tendo resolvido todas as pendências relativas à sentença anterior. Essa premissa revelou-se frágil, pois

surgiram evidências de que o intérprete, em muitas situações, segue adiante, incólume, com interpretações mais simples, distorcidas ou inconsistentes da sentença interpretada (FERREIRA; ÇOCAL, 2016).

Um exemplo de modelo que considera essas novas evidências está descrito num artigo de Ferreira e Patson (2007), chamado "*The 'Good Enough' Approach to Language Comprehension*". Em síntese, as autoras acreditam que o intérprete está empenhado em executar uma interpretação que é apenas boa o suficiente para satisfazer a tarefa que lhe foi imposta (por si, ou por outrem). A depender da tarefa, a interpretação pode ser mais ou menos completa, mais ou menos detalhada, e mais ou menos próxima das possibilidades oferecidas pela sentença interpretada.

Um exemplo fornecido pelas autoras é:

(10) The singer saw the audience member with the binoculars¹.

Que corresponde, sintaticamente, a uma sentença usada por Ribeiro (2015, p. 100):

(11) O policial viu o turista com o binóculo preto.

Em ambos os casos, as sentenças são ambíguas e o binóculo pode ligar-se tanto ao verbo, quanto ao sintagma nominal na posição de objeto direto, dando margem a duas possibilidades interpretativas (o turista foi visto com o uso de um binóculo vs. o turista que estava com um binóculo foi visto). Em muitos modelos, o intérprete sempre esgotaria seus recursos (informações disponíveis e capacidade de processamento) para chegar à interpretação mais provável. Segundo Ferreira e Patson (2007), no entanto, o empenho nesse esforço dependerá da tarefa imposta ao intérprete.

Elas propõem que a interpretação inicial é rápida e automática, e que qualquer esforço adicional só será empreendido em caso de necessidade. Elas comparam seu modelo ao modelo proposto por Tversky e Kahneman (1974), que analisaram, em situações não linguísticas, os vieses que influenciam escolhas feitas rapidamente em situações de incerteza. Como propõe Kahneman (2011) em seu modelo cognitivo para tomada de decisões em geral, nossa mente disporia de dois sistemas: o chamado Sistema 1, que operaria de forma rápida e automática, com pouco ou nenhum esforço e sem controle voluntário

¹ A cantora viu o espectador com o binóculo (tradução nossa).

perceptível; e o Sistema 2, que exigiria vontade, atenção e esforço, e corresponderia à nossa percepção subjetiva de que estamos realmente agindo, nos concentrando e fazendo escolhas.

No caso da sentença (11), segundo este modelo, o intérprete colocaria o binóculo nos olhos do policial ou do turista, de forma automática, a depender de uma série de fatores aos quais não prestaria muita atenção (Sistema 1). Uma análise mais profunda da ambiguidade da sentença (Sistema 2) só seria empreendida caso a tarefa exigisse.

2.4 Processamento de sentenças: alguns modelos emblemáticos

O campo de estudo do processamento de sentenças procura explicar, como vimos, a compreensão da linguagem por meio do estudo de suas fontes de informação, dos mecanismos de processamento dessas informações, e do grau de equivalência entre a sentença e a interpretação que o intérprete lhe dá. Antes de prosseguir com modelos específicos de processamento de sentenças isoladas, devemos lembrar, entretanto, que sentenças, via de regra, só costumam andar sozinhas em experimentos desenhados por linguistas. No mundo real, as sentenças costumam fazer parte de um conjunto maior, de uma superestrutura que lhes organiza e lhes dá sentido e que muitos chamam de discurso. Um modelo bastante aceito é o da estrutura atencional e intencional do discurso formulado por Grosz e Sidner (1986).

As autoras propõem-se a descrever como processamos o conjunto de enunciados de um determinado discurso; e dividem esse processamento em três níveis: um componente linguístico — o agrupamento de enunciados em segmentos do discurso —; um componente intencional — o reconhecimento das intenções de cada enunciado, de cada segmento e do discurso como um todo —; e um componente atencional — os mecanismos que gerenciam a atenção a cada trecho do discurso. Como resume uma das autoras — em Grosz, Weinstein e Joshi (1995) —, a coerência do discurso como um todo depende da coerência entre seus segmentos, e a finalidade de cada segmento serve à finalidade do discurso. Elas chamam seu modelo de "teoria computacional da estrutura do discurso".

A teoria é bastante abrangente e detalhada, mas, como bem aponta Payne (1997, p. 350), ao mencionar esquemas semelhantes, como a "*rethorical structure theory*" e a "*story grammar*", trata-se de abordagem mais próxima da interpretação do discurso do que, propriamente, da análise linguística. Existe certa subjetividade nas análises propostas; e abordagens de processamento da linguagem da mesma época, como a de Ferreira e Clifton Junior (1986), que será mencionada adiante, já

demonstravam que certas preferências de interpretação sintática locais independem de fatores semânticos, pragmáticos ou contextuais.

Além disso, deve-se notar que as “variáveis dependentes” neste estudo — as incongruências — serão obtidas pela observação da leitura em voz alta. A fala produzida por meio da leitura é diferente da fala espontânea. Dehaene (2009) demonstra que o cérebro humano tem duas rotas diferentes para o processamento da leitura. Após o processamento inicial em áreas visuais, há uma bifurcação condicionada pela tarefa imposta ao leitor: se a tarefa privilegia a compreensão do texto — ou, em termos mais específicos, o acesso ao significado das palavras —, a rota passa pelas regiões inferiores do lobo temporal; se a tarefa privilegia a produção correta dos sons da fala, a rota segue por regiões mais superiores do mesmo lobo. Esse tipo de dado neurobiológico é coerente com nossa percepção subjetiva de que há duas possibilidades para a leitura em voz alta: é possível ler compreendendo detalhadamente o texto; ou ler sem se preocupar com o sentido e as implicações do texto, impondo-se apenas a tarefa de reproduzir seus aspectos articulatórios e prosódicos.

Considerando isso, mas, ao mesmo tempo, reconhecendo a necessidade de analisar, pelo menos, a relação entre sentenças adjacentes, optamos por utilizar, em nosso estudo, a ideia dos “segmentos do discurso” (DS) proposta por Grosz e Sidner (1986) na página 183 de seu artigo. Nossa interpretação é de que a relação entre “segmentos do discurso” é mais próxima de nossa proposta, pois é que há de mais objetivo no modelo das autoras; e, ademais, no exemplo ilustrado no artigo, o modelo foi aplicado justamente a um texto escrito argumentativo semelhante aos textos que serão analisados em nosso estudo.

Descendo do nível do discurso para o nível das unidades de que se compõe — as sentenças —, revisaremos alguns modelos de processamento de sentença propostos nas últimas décadas. Tais modelos ainda servem, atualmente, como pressuposto teórico para novos estudos, e nossa revisão, novamente, se baseará na análise de Ferreira e Çokal (2016), com alguns acréscimos.

O mais antigo e, talvez, o mais influente desses modelos é o chamado ‘modelo de dois estágios’ (*two-stage model*), mais conhecido como ‘teoria do *garden-path*’ (TGP). Esse modelo considerava que o intérprete construía a estrutura da sentença apenas com base em regras sintáticas, sem acesso inicial ao léxico, e com as restrições impostas por duas operações fundamentais: o *late closure* — informação nova é ligada ao constituinte atual; e o *minimal attachment* — informação nova é ligada ao nó sintático

atual, evitando a criação de novos nós. Em PB, são utilizados, respectivamente, os termos *aposição local* e *aposição mínima* (MAIA; FINGER, 2005).

Declaração recente de Ribeiro (2015, p. 93), dizendo que a TGP é "ainda o modelo teórico de maior relevância no campo do Processamento de Frase, aos trinta e cinco anos do seu surgimento no cenário da Psicolinguística Experimental", parece indicar que essa teoria continua sendo bastante influente em nosso País. Outra evidência nesse sentido vem de Maia (2019, p. 564), que afirma, textualmente, que a área de estudo do processamento de sentenças:

"tem sido marcada pela oposição entre duas classes de modelos incompatíveis. De um lado, modelos baseados no paradigma simbólico, tal como a Teoria do *Garden Path* — TGP (...) —, que assume um processador em dois estágios (*parsing* e interpretação), propõem uma análise sintática inicial, serial e incrementacional, baseada em métricas de simplicidade, atrasando o acesso a informações semânticas e pragmáticas para a fase interpretativa, pós-sintática. De outro lado, modelos conexionistas, baseados no paradigma de redes (...), propõem que o processador ative em paralelo múltiplas fontes de informação que podem competir entre si, produzindo ativações com diferentes pesos".

Ferreira e Çocal (2016) prosseguem em sua revisão dizendo que nas últimas três décadas, com base em diversas evidências experimentais, sugeriram-se algumas modificações na TGP. A principal delas foi a noção de "*construal*", que talvez possa ser traduzida em PB como *interpretação*. A razão para essa nova abordagem é que, em algumas situações experimentais, o intérprete lê mais rápido sentenças ambíguas (p.ex.: "o empregado da atriz que estava na varanda") do que sentenças não ambíguas. Entretanto, se houver uma pergunta requerendo resolução da ambiguidade, a leitura dessas sentenças torna-se mais lenta do que a de suas contrapartes não ambíguas. Em contraste com a TGP — mas ainda prestando continência aos modelos sintaticistas —, o *construal* sugere que, durante a leitura, determinados constituintes novos poderiam não ser definitivamente ligados — mas apenas associados provisoriamente — à estrutura que já foi construída até aquele momento, o que seria feito de forma mais rápida do que uma ligação sintática definitiva.

Uma série de modelos recentes comunga esse pressuposto (e não muito mais do que isso) de que o intérprete nem sempre finaliza o processamento com uma interpretação fiel à sentença ouvida/lida; e que interpretações incompletas, ou mesmo deturpadas, são bastante comuns. Tais modelos consideram que o processamento das sentenças é muitas vezes superficial, o chamado *shallow*

processing, como é o caso do ‘*Good Enough*’ *Approach to Language Comprehension* (FERREIRA; PATSON, 2007). Outro acréscimo à TGP que não deve ser ignorado foi o reconhecimento da prosódia como um dos fatores que interferem no processamento inicial da sentença e que modulam o funcionamento do processador sintático — o famoso *parser* —, dispositivo teórico fundamental das teorias derivadas do TGP.

Mas nem só de TGP vivem os teóricos. Contrapondo-se a essa vertente — que privilegia um processamento serial, em dois estágios, com o estágio inicial predominantemente, ou exclusivamente, sintático/gramatical —, surgiram teorias negando o modularismo e propondo uma base conexionista. Essas teorias se apoiam na hipótese de que o processamento de sentenças ocorre em paralelo, não em série; e que, desde as fases mais precoces, diversas fontes de informação podem ser acessadas. Uma das fontes de inspiração dessas teorias foram as ideias de Pesetsky (1995), expostas em sua obra “*Zero Syntax*”, que sugeriram um foco maior no léxico como fonte não apenas de representação semântica, mas também de organização sintática.

2.5 Evidências do cérebro em funcionamento: eletrofisiologia e neuroimagem funcional

Tendo feito uma exposição geral sobre o campo de estudo do processamento de sentenças, encerramos o vínculo com o esquema revisional proposto por Ferreira e Çocal (2016) e passamos a dar destaque a algumas considerações, estudos, teorias e modelos que, por determinadas razões, nos parecem relevantes para os objetivos deste trabalho. Começamos pelas considerações de Bornkessel-Schlesewsky, Staub e Schlewsky (2016), que fazem uma análise crítica das principais ferramentas investigativas que nos permitem observar mais de perto o funcionamento do cérebro e que são cada vez mais utilizadas nos estudos sobre o processamento da linguagem.

Hoje em dia, diversos estudos empregam métodos eletrofisiológicos (como o *evoked related potential* — ERP) e de neuroimagem funcional (fMRI) e correlacionam os dados obtidos por esses exames neurofisiológicos com dados linguísticos escolhidos em desenhos experimentais. Esses métodos têm vantagens e desvantagens distintas. O ERP oferece ao investigador uma boa resolução temporal, na ordem dos milissegundos, mas possui resolução espacial ruim, na ordem dos centímetros, não sendo ideal para localizar espacialmente os fenômenos neurais. A fMRI, por sua vez, tem uma resolução espacial que pode ser muito boa, mas sua resolução temporal situa-se na casa das dezenas de segundos; ou seja, localiza no espaço, mas não localiza no tempo. Há, também, instrumentos com

excelente resolução temporal e espacial, como a eletrocorticografia (ECoG), que aplica eletrodos diretamente no tecido cerebral, mas que, justamente por ser tão invasivo, tem limitações óbvias de aplicabilidade no dia a dia, sendo usado apenas em condições excepcionais, como durante cirurgias para controle de epilepsia.

Por seus pontos fortes — excelente resolução temporal e relativa facilidade de acesso — e apesar de seu ponto fraco — péssima resolução espacial —, o ERP tem sido muito usado em estudos linguísticos. Seus resultados costumam ser ilustrados por meio um gráfico em que o eixo das abcissas (x) ilustra a passagem do tempo em milissegundos e o eixo das ordenadas (y) ilustra o potencial elétrico, positivo ou negativo, predominante naquele momento em determinada região do couro cabeludo (com a esperança de refletir a atividade cerebral subjacente). Segundo Friederici (2011), as correlações mais relevantes entre os eventos linguísticos e eletrofisiológicos observados no exame de ERP seriam as seguintes:

1. *Early left anterior negativity* (ELAN), que ocorre entre 120 e 200 milissegundos (ms) e é associado à construção inicial da estrutura sintática da sentença;
2. N400, que ocorre entre 300 e 500 ms e reflete processos semânticos;
3. P600, que ocorre por volta de 600 ms e reflete processos sintáticos tardios, como o processamento de anomalias sintáticas.
4. Outros fenômenos:
 - a. N100, ocorre por volta de 100 ms e é associado ao "processador acústico-fonético"
 - b. LAN, negatividade anterior esquerda, na mesma janela de tempo da N400, mas relacionada a relações gramaticais verbo/argumento

Apesar desses dados transmitirem uma impressão de universalidade e uma sensação de autoridade biológica, Bornkessel-Schlesewsky, Staub e Schlewsky (2016) sugerem que, em virtude das limitações inerentes aos métodos (ERP e fMRI), devemos ter cautela na correlação entre os achados neurofisiológicos (que não sabemos, realmente, de onde vêm e o que significam) e os fenômenos linguísticos que escolhemos analisar (e que são, na verdade, apenas construtos teóricos criados pelo pesquisador).

2.5.1 Das evidências neurobiológicas aos modelos cognitivos (e vice-versa): linhas tortuosas

Para mostrar que não há uma relação muito direta entre evidências neurobiológicas e modelos linguísticos cognitivos, Bornkessel-Schlesewsky, Staub e Schlewsky (2016) mencionam dois modelos antagônicos de processamento de sentenças que, baseados em ERP, interpretam resultados eletrofisiológicos idênticos de forma diferente: o "*neurocognitive model of auditory processing*", um modelo de base serial e modular, defendido e muito bem explicado por Friederici (2011); e o "*memory, unification and control (MUC)*", de base paralela e interacionista, de Peter Hagoort. Segundo Bornkessel-Schlesewsky, Staub e Schlewsky, a divergência na interpretação de dados eletrofisiológicos objetivos só se explica pelas diferentes filiações e pontos de partida teóricos dos defensores de cada modelo.

Além da heterogeneidade na interpretação dos dados obtidos por meio de ferramentas neurofisiológicas, estão surgindo evidências contra a universalidade dos mecanismos de processamento de elementos linguísticos que se julgava como idênticos em diferentes línguas. Têm-se observado achados neurofisiológicos distintos associados a tarefas linguísticas hipoteticamente similares em línguas diferentes, o que sugere que a forma de organização de determinada língua pode influenciar o processamento da linguagem em nível neurobiológico. Bornkessel-Schlesewsky et al. (2011), por exemplo, estudaram diversas línguas — inglês, holandês, islandês, alemão, turco e mandarim —, correlacionando os achados de ERP a um paradigma experimental chamado “reversão semântica”. O paradigma de reversão semântica expõe os indivíduos a sentenças do tipo "o anzol lançou o pescador", onde a sentença é sintaticamente bem-formada, mas semanticamente implausível. Em inglês e holandês, eles observaram, associado à reversão semântica, um padrão de P600; em turco e mandarim, N400; e em alemão, um padrão duplo, com N400 e positividade tardia. A conclusão foi que esses eventos neurofisiológicos não estão necessariamente relacionados de forma específica a construtos teóricos cognitivos (como, por exemplo, o processamento semântico ou sintático), mas a estratégias de mapeamento forma-significado específicas de cada língua, segundo sua organização morfossintática.

2.5.2 Modelos cognitivos: alguns problemas fundamentais

Como vimos, os modelos de processamento de sentenças divergem em praticamente todos os pontos em que poderiam divergir. Entretanto, costumam concordar com relação a alguns pressupostos

fundamentais, como, por exemplo: a) deve ser possível estabelecer um modelo de processamento que seja universal e se aplique a todas as línguas; b) processamento de sentenças equivale à compreensão da língua, em oposição à produção da língua.

Esses pressupostos, no entanto, também não estão imunes a críticas. Evans e Levinson (2009) — com base em evidências da linguística tipológica e descritiva — questionam o que chamam de "mito dos universais linguísticos" nos modelos cognitivos da linguagem. Segundo os autores, faria mais sentido pensar nos tais universais como exceção, e na diversidade como regra. As regularidades observadas nas diversas línguas não justificariam, em sua opinião, uma busca por mecanismos neurobiológicos (e modelos cognitivos) universais dedicados especificamente à linguagem. A busca, segundo eles, deveria ser uma busca por mecanismos que permitissem ao cérebro adaptar-se à variabilidade linguística e cultural. Em suas palavras (EVANS; LEVINSON, 2009, p. 447), "*cognition is less like the proverbial toolbox of ready-made tools than a machine tool, capable of manufacturing special tools for special jobs*"¹.

Outro ataque à visão tradicional sobre o processamento da linguagem vem de Pickering e Garrod (2013). Eles criticam a visão esquemática e modular que divide o processamento da linguagem entre produção e compreensão e propõem um mecanismo unificado. Em sua visão, produção e compreensão estariam interligados. O intérprete faria uso de mecanismos de produção para auxiliar na compreensão das sentenças e para se preparar — prever —, com base em sua própria experiência e capacidade de produção, as sentenças que viriam na sequência. Os autores usam como uma das justificativas para essa teoria o fenômeno de "*covert imitation*", em que um observador de qualquer tipo de ação (motora ou linguística) imita ou simula, de forma muitas vezes involuntária, a ação que está observando.

Essa visão unificada dos processos de produção e compreensão da língua guarda certa semelhança com as observações feitas por Patel (2006) em animais não humanos. Estudando a relação entre música e linguagem — mais especificamente, o fenômeno de "*beat perception and synchronization* (BPS)" —, Patel (2006, p. 101) diz que "*vocal learning requires a tight coupling*

¹ A cognição é menos uma caixa de ferramentas do que uma máquina capaz de produzir ferramentas específicas para trabalhos específicos. (tradução nossa)

between auditory input and motor output in order to match vocal production to a desired model"¹. Segundo essa hipótese, nossa capacidade de produzir movimentos e vocalizações sincronizadas com um ritmo externo (como na dança e no canto) estaria intrinsecamente vinculada à nossa capacidade linguística. Sendo assim, a BPS só poderia ser encontrada em espécies de animais não humanos que tivessem capacidade de aprendizagem vocal, como papagaios, corvos e cetáceos. E, de fato, em Patel (2009), descreve-se um espécime de cacatua (*Cacatua galerita Eleonora*, um tipo de papagaio) capaz de sincronizar seus movimentos com o ritmo de uma música.

2.6 A(s) memória(s) — limitações e consequências linguísticas

A aplicabilidade universal dos modelos cognitivos de processamento da linguagem e os reais mecanismos que regem esse processamento não estão, como vimos acima, muito bem estabelecidos. Ambos os problemas derivam da dificuldade para estabelecer modelos que reflitam a realidade do processamento de sentenças, uma vez que essa realidade está, em grande parte, enterrada nas profundezas do sistema nervoso. Os fenômenos linguísticos que observamos no dia a dia e que estão disponíveis para nosso estudo são, na verdade, reflexos de processos neurobiológicos que, em sua maior parte, são completamente inacessíveis à nossa investigação. Há, entretanto, determinados fenômenos neurobiológicos que, apesar de não estarem completamente desvendados ou expostos, podem ser considerados como realidades objetivas incontornáveis e, como tal, podem e devem servir de baliza a qualquer modelo cognitivo de processamento de linguagem que pretenda ancorar-se na realidade.

Grodner e Gibson (2005), por exemplo, analisaram a leitura de sentenças em inglês por meio da técnica de leitura automonitorada. Eles observaram que um fator independente de dificuldade de leitura é a quantidade de palavras entre a palavra atual (que está sendo lida) e a palavra ou constituinte com o qual a palavra atual estabelece relação de dependência. Por exemplo, há mais "dificuldade" (maior tempo de leitura) para a leitura do **verbo** em "*the nurse from the clinic supervised the doctor while...*"² do que em "*the nurse supervised the doctor while...*". Esse fenômeno, que a TGP chamou de "late closure" (aposição local) — a tendência, por exemplo, de ligar o verbo ao NP mais próximo

¹ O aprendizado vocal requer uma integração estreita entre o *input* auditivo e o *output* motor para que as produções vocais sejam coerentes com o modelo pretendido. (tradução nossa)

² A enfermeira da clínica supervisionou o médico enquanto... (tradução nossa)

(...*the clinic supervised...*) —, recebeu dos autores uma explicação com base em uma realidade neurobiológica bem estabelecida: a memória de trabalho.

De fato, não é possível compreender o processamento de sentenças sem considerar a memória de trabalho (WM, *working memory*). E, no caso do presente estudo, que lida com a leitura em voz alta, cujos pronunciamentos são lidos, e não recitados, justamente função de limitações de memória, torna-se ainda mais importante considerar a memória, *lato sensu*, como uma barreira cognitiva que precisa ser levada em conta tanto na análise das diversas teorias de processamento de sentenças, quanto na interpretação das incongruências encontradas. A WM, também conhecida como memória de curto prazo, foi definida por Baddeley (1992) como um sistema cerebral que permite o armazenamento temporário e a manipulação das informações necessárias à realização de tarefas tão complexas quanto o raciocínio, o aprendizado e a compreensão da linguagem. O modelo pioneiro proposto por Baddeley (1992) sofreu modificações e já não considera a WM como responsável pela “manipulação” de informações (CAPLAN, 2016). Dehaene (2014), entretanto, acredita que o escopo da WM pode ser ainda maior, estando intimamente ligada à própria consciência, mencionando quem acredite que a principal função da consciência seria a de transferir elementos da WM para a memória de longo prazo: ou seja, a consciência existiria a serviço da memória.

Caplan (2016) fez um apanhado das relações entre WM e compreensão de sentenças e propôs um modelo que pode ser útil ao nosso raciocínio. Segundo ele, A WM é capaz de manter sob sua guarda — de forma temporária, transitória e dependente de atenção (ou, conforme Dehaene (2014), de consciência), à disposição dos processos conscientes — algo entre **3 e 5 unidades de informação**, independentemente da modalidade (sons, imagens etc). Aplicando esse raciocínio a fenômenos linguísticos, parece natural concluir que a WM seja capaz de reter sob sua guarda de 3 a 5 palavras. Entretanto, é sempre útil lembrar que uma palavra, considerada como uma unidade de informação, é um complexo composto por diversas **subunidades de informação**. Cada palavra é um aglomerado de componentes fonéticos, sucessivamente combinados e interpretados como fonemas, que são combinados e interpretados como morfemas e, só então, combinados e interpretados como uma palavra, *stricto sensu*. Sendo assim, para que a WM seja capaz de reter de 3 a 5 palavras, as subunidades de informação de cada palavra (sons, fonemas, morfemas etc) deverão ter sido previamente processados, reinterpretados e agregados como unidades de informação de nível superior àquelas que lhes deram origem.

Por isso, Caplan (2016) apela para o conceito de "memória de trabalho de longo prazo" (LT-WM) (ERICSSON; KINTSCH, 1995, apud CAPLAN, 2016). A LT-WM foi um conceito criado para explicar a extraordinária capacidade de memória que certas pessoas — como, p. ex., jogadores de xadrez, calculistas e indivíduos que ganham a vida dando demonstrações de sua capacidade de decorar listas telefônicas, listas de palavras etc — desenvolvem em resposta a tarefas muito específicas. Ericsson e Kintsch (1995) propõem que essas habilidades, que eles chamam de "*expert memory*", são o resultado de atividades especializadas de domínio específico, não transferíveis para outros domínios de atividades não familiares. Ou seja, são limitadas às atividades especializadas em que se desenvolveu a LT-WM e decorrem da construção, por meio do aprendizado, de uma habilidade para agregar diversas unidades de informação em uma só unidade, com pistas específicas que permitem sua codificação e seu acesso na memória de longo prazo.

Caplan (2016), então, propõe que o processamento de sentenças seria uma atividade especializada decorrente de um tipo de "*expert memory*" e que se utilizaria largamente da LT-WM. O intérprete, dessa forma, recorreria a uma espécie de memória procedural, adquirida ao longo de anos e anos de prática com a língua. Ele seria, portanto, um *expert* na interpretação da própria língua, assim como um pianista tocando piano, um médico fazendo diagnósticos, ou um Grande Mestre jogando xadrez.

Christiansen e Chater (2015) trabalharam em cima dessas ideias e propuseram um modelo de processamento da linguagem a que denominaram "*now or never bottleneck*". Eles argumentam que a WM é uma limitação fundamental ao processamento da linguagem e que, para superar essa limitação, o cérebro teve de desenvolver estratégias muito específicas de processamento. Seu modelo é extenso e procura abarcar o fenômeno de maneira abrangente; mas, em resumo, eles propõem que a estrutura das línguas é fortemente influenciada pelas limitações de processamento, especialmente as limitações da WM. Durante o processamento de sentenças, as informações são recebidas sequencialmente e precisam ser rapidamente recodificadas para não se perderem devido ao fluxo ininterrupto de novas informações que continuam chegando. Essa recodificação seria feita por meio de um processo que eles chamam de "*chunk-and-pass*", onde, para manter o volume de unidades de informação dentro dos limites da WM, as informações de nível inferior seriam rapidamente processadas e agrupadas em níveis hierárquicos superiores, mais abstratos, aglutinando cada vez maior quantidade de informação em unidades autônomas e manipuláveis: da acústica para a fonologia, e daí para a palavra, para os constituintes, para a sentença, para o discurso e assim por diante.

2.6.1 O papel da antecipação no processamento de sentenças

Além desse processamento rápido, os autores também propõem um fenômeno de antecipação, de processamento preditivo, “adivinhandando” informações que ainda não chegaram. Com base nas informações que acabaram de ser processadas, haveria mecanismos probabilísticos capazes de antecipar palavras, estruturas sintáticas, conceitos, contextos discursivos etc., tornando o processamento ainda mais rápido e eficaz. Uma implicação desse modelo, segundo os autores, é que o aprendizado de uma língua não consistiria em configurar sua gramática, como muitos propõem, mas em aprender a processar seus elementos.

A rapidez no processamento das sentenças e a fluência na produção da fala, portanto, dependeriam de um acesso rápido a elementos armazenados na memória (de longo prazo, LTM). Elementos cognitivos mais salientes na memória, mais disponíveis ao acesso, teriam maior probabilidade de serem utilizados. Essa ideia deu impulso ao estudo do fenômeno de priming em linguística: quando o cérebro recebe um estímulo qualquer, a representação cerebral desse estímulo permanece saliente durante algum tempo, tornando mais provável que essa representação seja incorporada a processos cognitivos imediatamente subsequentes.

Caplan (2016), citando diversos autores, explica a importância do *priming* dizendo que, se há, em determinado momento, uma concorrência entre dois sinônimos igualmente adequados, o escolhido será aquele que estiver mais saliente na memória. Essa saliência seria possível não apenas para palavras, mas também para aglomerados (*chunks*) de palavras conectadas sintaticamente, o que daria origem ao *priming* estrutural (para estruturas sintáticas). O *priming* estrutural tornaria mais salientes na memória e mais disponíveis para uso imediato determinados arranjos sintáticos usados recentemente; e acredita-se que o fenômeno do alinhamento entre participantes de diálogos, quando os indivíduos parecem convergir em suas formas de falar, seja uma de suas consequências.

Jackendoff (2015) exemplifica o *priming* sintático demonstrando que, se o intérprete ouve uma sentença como (12) (que ele chama de "ditransitiva"), torna-se mais provável que ele produza, na sequência, uma sentença ditransitiva como (13), em vez de uma sentença como (14) (que ele chama de "dativa preposicional"):

(12) He gave Bill the ball.¹

(13) She sent Harry the package.

(14) She sent the package to Harry.

Mas a antecipação do que está por vir — considerada como um dos componentes que explicariam a eficiência do processamento de sentenças, em particular, ou da linguagem, em geral — não poderia apoiar-se exclusivamente no *priming*. Ferreira e Chantavarin (2018) dizem que, segundo diversos autores, o *priming* não poderia ser considerado como um exemplo de antecipação, de predição do que está por vir, pois seria um fenômeno neural passivo, primário; e a predição teria de ser ativa, construtiva, um processo do tipo "*top-down*"², não "*bottom-up*"³. As autoras lembram que as primeiras teorias de processamento de sentenças fixavam-se nos mecanismos “passivos” de incorporação sucessiva de novas informações às informações previamente incorporadas e que a possibilidade de predição era pouco considerada. Hoje em dia, a possibilidade de predição, principalmente lexical, tem sido bastante estudada, e os pesquisadores acreditam que os intérpretes podem inferir significados antes mesmo do input da informação nova.

As autoras propõem uma reconciliação entre essas duas abordagens sob a forma de um modelo em que as informações são integradas ao contexto linguístico e ao conhecimento prévio, gerando uma rica representação semântica e um estado de preparação cognitiva para ideias, consequências e, eventualmente, até palavras específicas que teriam maior probabilidade de surgir no discurso.

Pickering e Garrod (2013) também defendem um modelo ativo, bastante detalhado e complexo de predição. Eles extrapolam fenômenos observados no comportamento motor não linguístico — em que a percepção de uma ação externa desencadeia mecanismos de ação (submotora ou abertamente motora) por parte de quem percebe — para o campo linguístico. Eles propõem que a compreensão (da língua) envolve mecanismos de produção (da língua) e vice-versa. Produção e compreensão seriam

¹ (12) Ele deu a bola para o Bill. (13) Ela enviou para o Harry o pacote. (14) Ela enviou o pacote para o Harry. (tradução nossa)

² De cima para baixo (tradução nossa)

³ De baixo para cima (tradução nossa)

indissociáveis, ocorreriam sempre em conjunto, e serviriam uma à outra como mecanismos de controle e de predição de eventos linguísticos futuros.

2.7 *Chunking, entrenchment* e outros modelos cognitivos

Contornando as limitações da WM e explicando o processamento de sentenças e a aquisição das línguas

Outro modelo que procura integrar as diversas questões que dividem o campo do processamento de sentenças (incluindo as línguas de sinais) leva o nome de "*multiscale information transfer framework* (MSIT)" e foi proposto por Blumenthal-Dramé e Malaia (2018). Assim como Christiansen e Chater (2015), elas defendem que os sinais (sejam acústicos, como nas línguas orais; ou visuais, como em línguas de sinais) são processados hierarquicamente, em múltiplos níveis. E, assim como Pickering e Garrod (2013), elas consideram que existe uma significativa superposição dos circuitos cerebrais responsáveis pela compreensão e pela produção da linguagem, em todos os níveis hierárquicos propostos: da fonética à semântica, e vice-versa. E expandem seu modelo acreditando que ele que pode ser aplicado a mecanismos de processamento de informação que vão além da linguagem.

Um dos elementos mais importantes para viabilizar o modelo de Blumenthal-Dramé e Malaia (2018) é o conceito de "*chunk*", que já foi mencionado de passagem nessa revisão. As autoras explicam, em linhas gerais, que *chunk* equivale a um agrupamento de informações: informações isoladas são de alguma forma conectadas e, apesar de manterem sua autonomia como informações individuais, passam a compor um conjunto que, para efeito de processamento, equivale a uma única e nova informação. Ou seja, passa a ser possível acessar um conjunto de informações conectadas (de forma serial, por exemplo, como uma sequência de atos motores ou linguísticos) como se fossem uma só informação. Isso se aplica, por exemplo, a comportamentos motores automáticos ou semiautomáticos, como amarrar sapatos sem pensar em cada movimento isolado dos cadarços, ou compreender a palavra "linguagem" sem prestar atenção a seus componentes fonéticos, fonológicos ou, até mesmo, à palavra como um todo. A formação de *chunks* (chamada *chunking*) resulta, via de regra, de uma automatização via repetição, e está intrinsecamente ligada ao conceito de congruência, tanto na produção e compreensão da linguagem, quanto na realização e interpretação de atos motores variados.

O conceito de *chunking* está relacionado a outro conceito das ciências cognitivas, o chamado "*entrenchment*"¹. Esse conceito, formulado originalmente pelo linguista Ronald Langacker, é definido por Schmid (2017, p. 414) como sendo "*a set of cognitive processes — mainly memory consolidation, chunking, and automatization — taking place in the minds of individual speakers*"², ou, adicionalmente, os efeitos desses processos. Gobet (2017) elabora melhor esse conceito e diz que *entrenchment* se baseia na ideia de organização cognitiva, com agrupamento e reforço gradual de unidades. Acrescenta que o conceito tem sido proposto tanto como um mecanismo cognitivo geral, quanto como um mecanismo linguístico. E conclui afirmando que os modelos que levam em conta esse fenômeno costumam se opor aos modelos inatistas de aquisição de linguagem, pois propõem que a aquisição da linguagem é um processo cognitivamente análogo à aquisição de *expertise* em um domínio cognitivo qualquer. Neste trabalho, consideraremos *entrenchment* como sinônimo de *chunking*, pois, como vimos, o termo inclui consolidação de memória, automatização, organização cognitiva e agrupamento de unidades.

Um aspecto bastante interessante desse fenômeno — que talvez demonstre que o *chunking* é uma tendência cognitiva — é observado nas "*holophrastic units*", frases que são pronunciadas como se fossem uma só palavra, como se observa durante a aquisição da linguagem pelas crianças. Schmid (2017) fornece, como exemplo, "*more milk*"³, e acrescenta que a expressão não resulta de um processo de *chunking* de duas palavras, pois as crianças não identificam, inicialmente, cada palavra isoladamente, mas do *chunking* de uma série de fonemas. Ou seja, a sequência de fonemas que resulta em "*moremilk*" seria aprendido pelas crianças como se fossem uma unidade (uma palavra) e só mais tarde seria decomposta e utilizada em operações do tipo "*more X*" (*more tea, more toast, etc.*)⁴.

Em PB, a questão foi estudada por Abaurre e Silva (1993), analisando o desenvolvimento dos critérios de segmentação da escrita em crianças. As autoras não consideraram o conceito de *chunking*, mas ofereceram uma série de exemplos de casos comuns de hipossegmentação da escrita em PB. Como

¹ A literatura em português tem feito traduções diferentes do termo *entrenchment*. Neste texto, vamos considerar que a palavra se traduz como "entranhamento", e que o verbo *to entrench* traduz-se como "entranhar".

² Um conjunto de processos cognitivos — principalmente consolidação da memória, *chunking* e automatização — que acontecem nas mentes dos falantes (tradução nossa)

³ Mais leite (tradução nossa)

⁴ (mais chá, mais torrada, etc.) (tradução nossa)

hipóteses não mutuamente excludentes para a hipossegmentação, as autoras sugerem aspectos prosódicos e semânticos. Alguns exemplos oferecidos, extraídos de "um corpus constituído por centenas de textos espontâneos representativos da escrita inicial de crianças brasileiras (de escolas públicas e particulares, da pré-escola e da primeira série, de diversas regiões do país e de diferentes classes socioeconômicas)", seriam: omebateu (homem bateu), pediarvere (pé de árvore), pedefegão (pé de feijão), amigodoleão, cachorociente (cachorro quente), probepato (pobre pato), eufui (eu fui), lavois (ela foi), élagosta (ela gosta), iliqueria (ele queria), elecaiu (ele caiu), minhatia (minha tia), deouro (de ouro), umdia (um dia), apata (a pata), cesauvar (se salvar), oqueijo (o queijo), aminhacaza (a minha casa), votelefona (vou telefonar), todumundo (todo mundo), elaficoucotete (ela ficou contente) etc. A esses exemplos, poderíamos acrescentar um caso recente visto na internet, escrito por um menino de cinco anos: "éverdade esse bilete". Não vamos entrar na discussão dos possíveis mecanismos cognitivos por trás da hipossegmentação da escrita, mas citamos o fenômeno por considerá-lo interessante do ponto de vista do processamento da linguagem e da relação entre fala e escrita.

Stefanowitsch and Flach (2017) abordam outra questão relativa ao *chunking*, que é a pesquisa do fenômeno em análise de corpus. Inicialmente, eles defendem esse tipo de estudo, argumentando que, apesar da atitude hostil de certas facções linguísticas com relação ao estudo de corpus (*langue versus parole*, competência *versus* desempenho, *i-language versus e-language*), a incompatibilidade entre corpus e cognição seria apenas aparente. Eles citam duas visões que seriam favoráveis à utilização de corpora para os estudos relacionados à cognição linguística. A primeira, mais evidente, seria a visão de "corpus-as-output". Nessa perspectiva, um corpus seria uma amostra delimitada de produção linguística que refletiria mecanismos cognitivos. A segunda, menos usual, seria de "corpus-as-input", onde um corpus poderia, eventualmente, ser fonte importante de input linguístico, influenciando o uso da língua pelos que estão expostos a ele.

Depois da defesa inicial, os autores analisam a interpretação dos dados de frequência extraídos de um corpus. Eles advogam que os dados de frequência deveriam ser obtidos e interpretados segundo dois eixos ortogonais: complexidade e esquematismo. No eixo da complexidade, estariam sequencialmente classificadas: palavras constituídas por um morfema (desconsiderando vogais temáticas: dia, ouro etc.); palavras flexionadas e derivadas (diurno, dourado, douradas); e expressões fixas (boa tarde, partido político). No eixo do esquematismo, teríamos, sequencialmente: palavras concretas; e esquemas (palavras gramaticais, afixos e estruturas sintáticas do tipo NP V NP). A justificativa para essa divisão, segundo os autores, é que qualquer uma dessas realidades linguísticas

(estruturas concretas e estruturas esquemáticas) distribuídas ao longo desses dois eixos, de palavras a esquemas sintáticos, poderia ser objeto de *chunking*.

Para dar prosseguimento a essa revisão sobre WM e *chunking*, tomaremos emprestadas duas sentenças utilizadas por Jackendoff (2015, p. 199). Em um artigo em que defende as teorias racionalistas de funcionamento cognitivo e compara a gramática gerativa com seu próprio modelo, o "*parallel architecture*" (PA), Jackendoff utiliza duas estruturas sintáticas: uma comum, do tipo nome e verbo (15), e outra idiomática, a que chama de "*correlative comparative*" (16). Seu objetivo é mostrar que, nessas sentenças, o intérprete deve esperar sete palavras antes de decidir qual estrutura sintática está sendo utilizada e que, apesar dessa espera, não parece haver qualquer sensação subjetiva de dificuldade na tarefa:

(15) *The more thorough but less immediately attractive theory often doesn't win.*¹

(16) *The more thorough but less immediately attractive a theory is, the worse its chances are.*

As palavras sublinhadas em (15) e (16) indicam o momento em que o intérprete pode, finalmente, compreender de que tipo de estrutura sintática se trata. E, apesar da espera de sete palavras antes da decisão pela estrutura do constituinte, não há, como diz Jackendoff, sensação subjetiva de dificuldade. Considerando as limitações da WM e os modelos discutidos até aqui, é possível imaginar que essas sete palavras vão sendo processadas e incorporadas a uma unidade semântica hierarquicamente superior, como na ideia de *chunk-and-pass* de Christiansen e Chater (2015). Talvez seja útil pensar no que diferencia (do ponto de vista do processamento de sentenças pelo intérprete) a sequência de palavras em (15) da sequência de palavras abaixo (criada por nós e obtida a partir do embaralhamento das palavras de (15)):

(17) *The but more win attractive often immediately theory thorough doesn't less.*²

¹ (15) A teoria mais detalhada, mas menos atraente à primeira vista, geralmente não vence. (16) Quanto mais detalhada e menos atraente à primeira vista é uma teoria, piores são suas chances. (tradução nossa)

² A mas mais vence atraente geralmente à primeira vista teoria detalhada não menos. (tradução nossa)

Naturalmente, é possível extrair um significado global de (15) e apenas fragmentos desconexos de (17), mas o que passa despercebido, via de regra, são os mecanismos que empregamos, involuntariamente, para agrupar as informações e extrair esses significados. Segundo o modelo de transferência de informação multiescala de Blumenthal-Dramé e Malaia (2018), existe uma hierarquia linguística com graus regressivos de granularidade de informação (níveis hierarquicamente mais altos seriam menos granulares) percorrendo o caminho fonema, sílaba, morfema, lexema, frase, sentença, enunciado e discurso. Ao longo dessa escala, as informações mais granulares (mais numerosas) são aglomeradas em unidades de hierarquia superior, com menor nível de granularidade e maior nível de abstração. Em (17), o mecanismo de *chunk-and-pass* pode ser aplicado do nível fonético até o nível lexical. Em (15), ele segue sendo aplicado a níveis "superiores" de representação, formando constituintes sintáticos e prosódicos, argumentos do verbo e, finalmente, a sentença como um todo, que representaria uma **unidade de informação** de nível semântico, ou discursivo, ou pragmático; ou que, para fins de simplificação, poderia ser chamado simplesmente de **mensagem**.

2.8 Constituintes sintáticos, estrutura argumental, argumentos, papéis temáticos, relações gramaticais e outras abstrações correlatas

Nesta seção, discutiremos (de maneira superficial e aplicada aos nossos objetivos) a existência e a função dos diversos constituintes propostos teoricamente, especialmente os relacionados à estrutura argumental das sentenças, e a possível utilidade prática de uma classificação semântica e/ou morfossintática dos constituintes/argumentos.

Carnie (2013), a respeito do que ele chama de "realidade psicológica dos constituintes" (ou, ainda melhor, da "constituibilidade" — *psychological reality of constituency*, no original), justifica a existência real desse conceito ("constituente") citando um experimento da literatura que demonstra que um som de "clique" produzido no meio de uma palavra em determinada sentença é percebido pelos intérpretes em posição diferente da realidade, sendo deslocado perceptualmente de sua posição original para uma fronteira entre dois constituintes. Nas sentenças abaixo, as chaves marcam os constituintes sintáticos (segundo Carnie (2013)), a barra inclinada ("/") marca a posição real do "clique", e o hashtag ("#") marca a percepção do clique pelo intérprete:

(18) *[In her hope of marrying] # An/na was impractical.*¹

(19) *[Harry's hope of marrying An/na] # was impractical.*

Apesar de o autor partir de uma perspectiva gerativista da linguagem, como ele mesmo reconhece — "*my research revolves around various topics within the Minimalist Principles and Parameters approach to syntactic theory*"² (CARNIE, 2019) —, nada impede que apliquemos o conceito de "constituency" de forma apartidária. A definição de "constituente" formulada por Trask (1996, p. 57) é "[*constituent is*] any part of a sentence which is regarded as forming a distinct syntactic unit within the overall structure of the sentence, on the ground that it behaves as a unit with respect to certain criteria, such as displacement, coordination, ellipsis and the possibility of its serving as antecedent to a pro-form. In a tree diagram, a constituent is represented as a branch dominated by a single node. Usually only a continuous sequence can qualify as a constituent, but some approaches permit the recognition at discontinuous constituents"³. A definição do próprio Carnie (2016) é ainda mais simples: sentenças são compostas por constituintes; e constituintes são grupos de palavras que funcionam juntas como uma unidade.

A definição de "argumento", por sua vez, ainda segundo Trask (1996, p. 20), seria "*A noun phrase bearing a specific grammatical or semantic relation to a verb and whose overt or implied presence is required for well-formedness in structures containing that verb. Arguments may be identified either in terms of grammatical relations (Subject, Direct Object, etc.) or in terms of semantic roles (Agent, Patient, etc.)*"⁴. Dessa forma, consideraremos "argumento" como sendo um tipo de

¹ (18) [Na esperança de se casar]#An/na era irrealista. (19) [A esperança de Harry de se casar com An/na] era irrealista. (tradução nossa)

² Minha pesquisa lida com diversos tópicos relacionados aos Princípios e Parâmetros da teoria sintática Minimalista (tradução nossa)

³ Constituinte é qualquer segmento de um período que possa ser apontada como uma unidade sintática distinta dentro da estrutura geral desse período, considerando que esse segmento se comporte como uma unidade frente a certos critérios, como deslocamento, coordenação, elipse, ou a possibilidade de servir como antecedente de uma pró-forma. Em diagramas de árvore, um constituinte é representado por um ramo dominado por um único nó. De um modo geral, os constituintes são formados por sequências de palavras adjacentes, mas algumas abordagens reconhecem a existência de constituintes descontínuos. (tradução nossa)

⁴ Um sintagma nominal que possui, com um verbo, uma relação gramatical ou semântica específica, e cuja presença, expressa ou implícita, é necessária para que as estruturas de que o verbo faz parte sejam bem formadas. Argumentos podem ser identificados por suas relações gramaticais (sujeito, objeto direto, etc.), ou por seus papéis semânticos (agente, paciente, etc.). (tradução nossa)

"constituente" e não aprofundaremos, neste momento, a revisão desse ponto específico. A revisão que levaremos adiante diz respeito a duas questões: a importância dos argumentos para a construção das sentenças; e a organização da estrutura argumental dos enunciados. Antes, entretanto, abordaremos a noção de causalidade segundo George Lakoff.

Lakoff (2012, p. 69) acredita que experimentamos a noção de causalidade (relação de causa e efeito) como uma espécie de gestalt (uma sensação/percepção/cognição única, global). A causalidade não seria um conceito semântico primitivo, básico, indecomponível, mas uma expansão metafórica do paradigma (protótipo) da manipulação direta de um paciente por um agente, que incluiria as seguintes características: *"the agent has as a goal some change of state in the patient; the change of state is physical; the agent has a "plan" for carrying out this goal; the plan requires the agent's use of a motor program; the agent is in control of that motor program; the agent is primarily responsible for carrying out the plan; the agent is the energy source (i.e., the agent is directing his energies toward the patient), and the patient is the energy goal (i.e., the change in the patient is due to an external source of energy); the agent touches the patient either with his body or an instrument (i.e., there is a spatiotemporal overlap between what the agent does and the change in the patient); the agent successfully carries out the plan; the change in the patient is perceptible.; the agent monitors the change in the patient through sensory perception"*¹. Por exemplo, nas sentenças “o macaco subiu na árvore” e “João subiu na vida” haveria uma expansão metafórica, uma extrapolação da gestalt do campo de aplicação da primeira sentença para o da segunda sentença, por meio de determinados aspectos comuns entre esses campos.

A manipulação direta de objetos concretos do mundo (ou sua observação) — tão comum em nosso cotidiano, desde os primeiros meses de vida — resultaria em um protótipo cuja expansão metafórica daria origem à noção de causalidade, experimentada como sensação gestáltica que se aplicaria a uma enorme variedade de situações práticas e linguísticas. É possível correlacionar essa visão com a noção de predicado e predicação. Recorrendo novamente a Trask (1996, p. 2013),

¹ O agente tem por objetivo mudar o estado do paciente; essa mudança de estado é física; o agente tem um plano para atingir esse objetivo; o plano requer que o agente use um programa motor; o agente tem o controle sobre esse programa motor; o agente é o principal responsável por executar o plano; o agente é a origem de energia (ou seja, o agente está direcionando sua energia para o paciente), e o paciente é o destino da energia (ou seja, mudanças de estado do paciente ocorrerão por causa dessa energia); o agente toca o paciente com alguma parte de seu corpo ou com algum instrumento (ou seja, há uma superposição no espaço e no tempo entre o que o agente faz e o que acontece com o paciente); o agente executa seu plano com sucesso; a mudança no paciente é perceptível; o agente monitora, por meio de uma percepção sensorial, as mudanças efetuadas no paciente. (tradução nossa)

predicado é definido como: a) uma oração que se liga a uma "noun phrase"¹ (NP) para constituir uma sentença; ou b) um verbo (ou estrutura verbal) que deve ligar-se a determinados argumentos para dar origem a uma sentença. Consideraremos, em nosso trabalho, apenas a definição em "b".

A definição de predicado (V) como um verbo (*lato sensu*) cuja ligação a determinados argumentos (A) forma uma sentença (S) torna a noção de sentença bastante simples, podendo ser hipoteticamente formulada como: S=V+A. Para incluir as sentenças com verbos sem argumentos, poderíamos formular que S=V(+A), sendo que os parêntesis indicariam opcionalidade. Nessa formulação, o predicado (V) seria o elemento fundamental de qualquer sentença, acima de seus próprios argumentos. Na verdade, essa questão não está pacificada.

Platão — (apud Izre'el, 2018, p. 1677) — teria dito que: "o discurso nunca é composto por nomes falados em sucessão, nem por verbos falados sem nomes". Izre'el (2018) discorda. Ele propõe que a predicação explícita, gramaticalizada, é desnecessária em muitos contextos discursivos, sendo apenas necessário haver algo que ele chama de "domínio predicativo". Ou seja, considerando a língua real do dia a dia, falada de forma integrada (incluindo a sintaxe, a prosódia, a estrutura do discurso e a estrutura informacional do contexto), o domínio predicativo eximiria um enunciado (unidade do discurso delimitada por fronteiras prosódicas) de conter predicação explícita. O domínio predicativo seria, então, a unidade essencial do enunciado e não seria, necessariamente, representado por um verbo. Ele dá exemplos do que denomina 'sentenças unipartidas' ("*unipartite*", no original em inglês), sentenças compostas por apenas um elemento explícito (verbo ou argumento). Reproduzimos abaixo um diálogo em que os participantes discutem se determinado tipo de cavalo — visto por um deles em uma viagem à Mongólia — seria realmente um cavalo. As sentenças foram traduzidas do hebraico para o inglês pelo autor, e do inglês para o PB por nós (as barras inclinadas duplas indicam fronteiras prosódicas terminais, identificando cada sentença/enunciado):

(20) — Cavalo real?//

(21) — Cavalo cavalo// Só mais curto...// As pernas encurtadas, tipo isso.//

Na mesma linha de raciocínio, podemos citar um capítulo de Gil (2014), no contexto de um livro em que diversos autores exploram os limites da complexidade gramatical. O autor analisa o

¹ Sintagma nominal

desenvolvimento da predicação em línguas crioulas e em línguas de sinais — consideradas ambas como línguas relativamente novas, recentes, as primeiras por definição, as segundas por informação histórica. A conclusão é que, via de regra, a predicação ainda é pouco gramaticalizada nessas línguas. O autor avalia duas propriedades gramaticais importantes — a marcação dos argumentos e a marcação de tempo/aspecto/moço (TAM) — e conclui que essas propriedades estão interligadas e, em última instância, refletem uma característica tipológica fundamental das línguas: o grau com que cada uma gramaticaliza suas referências à predicação. Línguas com argumentos (agente, paciente; sujeito, objeto) e TAM bem marcados gramaticalizaram fortemente sua predicação. Línguas em que essas marcações estão ausentes, ou são opcionais, não gramaticalizaram significativamente sua predicação. O argumento principal é que, independentemente do grau de gramaticalização, toda e qualquer língua mantém relativamente intacta sua capacidade predicativa.

Em resumo, a **predicação**, considerada como um enunciado (unidade do discurso delimitada por fronteiras prosódicas) autônomo (segundo Izre'el, 2018), parece não exigir, do ponto de vista comunicativo, a presença de uma estrutura gramaticalizada, uma estrutura representada, obrigatoriamente, por um verbo e seu(s) argumento(s).

Outro conceito que está intimamente ligado à predicação é a **transitividade**. De acordo com Hopper e Thompson (1980), a transitividade é uma propriedade geral das sentenças e reflete uma transferência de ação de um participante para outro. Ela se distribui de forma relativamente contínua ao longo de uma “escala de transitividade” que indica maior ou menor transitividade e que, segundo tradução de Cunha, Oliveira e Martelotta (2015) para os termos de Hopper e Thompson (1980), inclui elementos como: número de participantes; grau de cinesia; aspecto mais ou menos perfectivo e mais ou menos punctual do verbo; intencionalidade e agentividade do ator; afetamento e individuação do objeto; e polaridade e modalidade da oração. As inúmeras possibilidades de combinação dos diversos aspectos da escala de transitividade materializam-se sob formas gramaticais e discursivas variadas nas diversas línguas e servem tanto para diferenciar construções transitivas de intransitivas, como para identificar o papel de cada argumento nas construções transitivas, conforme sejam mais ou menos agentivos.

No que se refere a este estudo, parece útil compreender a transitividade como um fenômeno onipresente e provavelmente universal, e que se correlaciona com uma diversidade de papéis temáticos (agente, paciente, causa, instrumento, tema, experienciador, beneficiário, objetivo, locativo, alvo, fonte, etc., etc., etc.) que poderiam ser desempenhados pelos diversos argumentos. Não há dúvida de

que cada uma dessas características semânticas assumidas pelos argumentos e cada uma das relações que eles estabelecem — tanto entre si, quanto com o contexto em que se inserem (sentença, enunciado, discurso) — é, em algum sentido, relevante. Se não fossem, não seria possível nem mesmo identificá-las. Mas, do ponto de vista do processamento de sentenças, alinharemos-nos a abordagens mais reducionistas, em sentido amplo, como, por exemplo, a proposta por Dowty (1991).

Dowty (1991) não nega a existência de todo o espectro de papéis/relações temáticas que poderiam ser assumidos pelos argumentos, mas propõe agrupá-los em apenas dois "**protopapéis**" temáticos ("*thematic proto-roles*", no original): **agente e paciente**. Ele baseia essa proposta no problema da seleção de argumentos de um verbo transitivo. Em geral, serão selecionados dois argumentos e cada um deles deverá assumir um papel prototípico, independentemente das sutilezas semânticas que se lhes possam atribuir. Havendo apenas dois protótipos para escolher — agente e paciente —, a tarefa se tornaria mais simples. Uma hipótese adicional, e bem interessante, lançada por Dowty (1991) é que a existência desses protótipos tornaria a aquisição da linguagem pelas crianças mais fácil, pois, havendo um esquema agente/paciente em que basear as expectativas, tornar-se-ia bastante reduzida a incerteza quanto aos papéis de constituintes desconhecidos, que poderão e deverão ser agrupados em apenas duas categorias.

Bickel (2010), revisando a extensa literatura tipológica sobre as relações gramaticais (GR; "*grammatical relation*", no original), reforça a diferença que existe entre as GR e os papéis semânticos ("*semantic roles*", no original), também chamados papéis temáticos ("*thematic roles*", ou "*theta roles*", no original). GR equivaleriam a propriedades morfossintáticas relacionando um argumento a uma sentença. Papel semântico/temático/teta, por outro lado, relacionaria argumentos a verbos (ou predicados). Por mais que essa distinção categórica seja importante para a literatura tipológica e seja motivo de discussão entre as diversas escolas linguísticas, é interessante discutir propostas que reinterpretam esses conceitos analíticos sob o ponto de vista do processamento da linguagem. Como veremos, o intérprete, no calor do momento, tendo de processar um rápido fluxo sequencial de informações que lhe chega aos ouvidos ou aos olhos, é obrigado a lidar com todas as informações — morfossintáticas e/ou semânticas, além de pragmáticas — de forma integrada, simultaneamente, sem tempo para considerações analíticas.

Uma das características que diferenciam a espécie humana de outras espécies animais, além da linguagem, é o que se costuma chamar de teoria da mente — ToM —, ou mentalização ("*theory of mind*", "*mentalizing*", no original). Segundo Frith e Frith (2003), trata-se da capacidade de reconhecer

nos outros, em particular, e no mundo, em geral, a partir do conhecimento ou das percepções que temos a nosso respeito, uma projeção de nós mesmos. Isso permite entender e prever o comportamento de outras pessoas e, por extensão — apropriada ou não —, de seres não humanos. É uma característica que pode estar presente desde os primeiros meses de vida, mas que se manifesta de forma mais evidente a partir dos 18 meses. A ToM tem relação com a nossa habilidade de projetar a noção de "agency" (capacidade de agir de forma independente, segundo as próprias vontades) em outros seres. Frith e Frith (2003), a partir de estudos com ressonância magnética funcional, localizaram essa funcionalidade mental na região do sulco temporal superior (STS) do cérebro humano.

Trazendo esse conhecimento para a linguística, Bornkessel-Schlesewsky e Schlewsky (2009) identificaram que essa mesma região, em sua porção mais posterior (pSTS — sulco temporal superior posterior), é ativada em tarefas que requerem avaliação do grau de prototipicidade do papel desempenhado pelos argumentos de uma sentença durante sua interpretação. Esses autores elaboraram um modelo de processamento da linguagem chamado "*extended Argument Dependency Model*" (eADM), que se baseia em avaliações neurofisiológicas (fMRI e ERP) aplicadas ao processamento de sentenças em diversas línguas. Eles analisaram a interpretação online de sentenças transitivas (com dois argumentos, em regra) em línguas com tipologias variadas de gramaticalização de sua predicação.

O modelo (eADM) e os estudos associados apontam questões interessantes, que abordamos a seguir. O primeiro é que eles avaliaram diversas línguas tipologicamente diferentes, incluindo línguas com verbo final, pós-argumentos (SOV), como alemão e mandarim. Essas línguas permitem observar o processamento dos papéis dos argumentos antes de o intérprete alcançar o verbo. Um achado comum que eles observaram entre as diversas línguas, independentemente da posição do verbo, foi a presença da onda N400 no ERP em situações que eles chamam de dissonância da prototipicidade de papéis ("*role prototypicality mismatch*", no original). O N400 é uma onda negativa que surge na região centro-parietal (no "alto" da cabeça), cerca de 400 ms após o evento linguístico estudado, como foi mencionado anteriormente. Uma situação em mandarim que desencadeia N400, segundo os autores, é a seguinte:

(22) a. 王子被挑战者刺死了。

prince BEI contender stab-PFV

*'The prince was stabbed by the contender.'*¹

b. 王子被绳子勒死了。

prince BEI cord strangle-PFV

'The prince was strangled by the cord.'

As chamadas 'construções com BEI' são tidas como equivalentes da passiva. O argumento que vem após o "BEI" é, morfossintaticamente, o agente. Quando o intérprete alcança esse argumento e encontra, em "b", o argumento "cord", morfossintaticamente agente, mas semanticamente menos "agentivo" do que "prince", essa incongruência prototípica desencadeia o N400.

Em seu modelo eADM, Bornkessel-Schlesewsky e Schlewsky propõem uma alternativa aos modelos mais comuns, que consideram que a identificação do papel temático de um argumento seria feita pela via morfossintática, enquanto a identificação de sua prototipicidade semântica (+/- agente, +/- paciente), pela via semântica/pragmática, numa divisão de trabalho entre morfossintaxe (gramática) e semântica/pragmática. Eles chamam de "proeminência" o conjunto de características — tanto morfossintáticas (marcação de caso e ordem de constituintes), quanto semânticas (animacidade, definitude, pessoa) — que indicam o papel do ator/agente ("actor", no original), em oposição ao atuado/paciente ("undergoer", no original). E propõem o conceito de uma interface ao mesmo tempo morfossintática e semântica que seria responsável por computar a proeminência dos argumentos com base em escalas ordinais de elementos como:

- a. Marcação morfológica de caso (nominativo > acusativo; ergativo > absoluto)
- b. Ordem dos argumentos (argumento 1 > argumento 2)
- c. Animacidade (+ animado > - animado)
- d. Definitude/especificidade (+definido/+específico > -definido/-específico)
- e. Pessoa (primeira/segunda pessoas > terceira pessoa)

Nesse modelo, as línguas difeririam no grau em que gramaticalizariam a proeminência de seus argumentos, havendo línguas em que aspectos como "a" e "b" seriam mais estritos, e outras em que

¹ (22) a. O príncipe foi esfaqueado pelo adversário. B. O príncipe foi estrangulado pela corda. (tradução nossa)

"c" e "d" teriam mais peso, mas a computação da proeminência seria feita na mesma etapa, considerando elementos tanto morfossintáticos quanto semânticos.

Em revisão atualizada deste modelo, Bornkessel-Schlesewsky e Schlewewski (2014) argumentam que, apesar da enorme variedade tipológica das línguas conhecidas, é possível propor uma estratégia comum a todas elas no que se refere ao processamento de sentenças. Eles estudaram línguas como checheno, inglês, finlandês, alemão, hindi, islandês, italiano, japonês, mandarim, russo, tâmil e turco e encontraram evidências neurofisiológicas de uma estratégia comum de interpretação incremental a que denominaram "*actor identification strategy*"¹ (AIS).

Essa estratégia tem as seguintes características: o intérprete identifica o ator da forma mais rápida e precisa possível; todos os argumentos concorrem para assumir esse papel; e o intérprete favorece atores que ocupam posição inicial nas sentenças. Considerando sentenças transitivas com dois argumentos, o protótipo de ator incluiria as seguintes características: é aquele que controla a situação descrita pelo verbo; é quem percebe, sente, experimenta; é quem faz contato; é quem move; é quem é ativo; é quem possui.

Com relação à escala de proeminência, eles atualizaram a que foi citada no estudo anterior para aplicá-la apenas à identificação do ator, ignorando o atuado/paciente/*undergoer*, que seria, agora, definido e identificado por exclusão. O ator prototípico possuiria todas as características a seguir, enquanto atores menos prototípicos deixariam de possuir alguma(s) delas: + animado, + humano, + definido, + 1ª pessoa, + nominativo (em línguas nominativo-acusativas), + primeira posição na sentença.

A justificativa e a vantagem desse modelo, segundo os autores, é que ele permitiria um processamento incremental, rápido e eficiente das estruturas linguísticas para identificar o responsável pelos eventos considerados (o ator) e, assim, estabelecer a estrutura argumental e semântica da sentença. Se pensarmos nas hipóteses de Dowty (1991), mencionadas acima, que propõe protopapéis temáticos (agente e paciente), esta proposta de Bornkessel-Schlesewsky e Schlewewski (2014) parece ainda mais interessante do ponto de vista da eficiência e rapidez de processamento, tanto para o intérprete adulto, como para uma criança em fase de aquisição da língua.

¹ Estratégia de identificação do ator (tradução nossa)

2.9 A Prosódia (e outros fenômenos suprasegmentares)

Lobato (1986), com relação às sentenças de (4) a (7) (mencionadas anteriormente e reproduzidas a seguir), diz que a aceitabilidade de (7) é rejeitada unanimemente, e a de (6), quase unanimemente, apesar de todas elas serem gramaticalmente idênticas e igualmente corretas:

(4) O rapaz saiu.

(5) O rapaz que o homem viu sai.


(6) O rapaz que o homem que a moça convidou viu saiu.

(7) O rapaz que o homem que a moça que João beijou convidou viu saiu.

Ela acrescenta que o usuário da língua, em seu desempenho (no sentido gerativista, confrontado com competência), não apela para as regras gramaticais quando emprega estruturas como "boa noite", "como vai?", e outras "frases feitas". No caso das sentenças de (4) a (7), segundo a autora, ele apela para as tais regras, mas em (7) — e, para a maioria dos falantes, em (6) —, a mente humana não seria capaz de armazenar as etapas gramaticais necessárias ao processamento das sentenças. Lobato emprega, então, um esquema gráfico, ligando o verbo a seu argumento, e faz uma analogia com os auxílios gráficos usados em cálculos aritméticos, em que é frequentemente necessário organizar graficamente o raciocínio, apesar de se conhecerem as regras aplicáveis. Com esse auxílio gráfico, a sentença (6) se tornaria inteligível para a maioria das pessoas, e a (7), para algumas:

(7)

O rapaz que o homem que a moça que João beijou convidou viu saiu.



Lobato (1986) lembra que certos recursos linguísticos desempenham a mesma função que o esquema gráfico acima e cita, como exemplos, os "elos semânticos" e a "entoação". Em última análise, trata-se da mesma questão abordada pelos modelos que citamos acima: como o intérprete organiza a estrutura argumental da sentença? Mas é interessante notar que os modelos que mencionamos não citam a prosódia como um elemento significativo do processamento inicial das sentenças. E, na verdade, não a incluem em nenhuma etapa do processo. No entanto, há estudos que tentam desvendar essa relação entre prosódia e processamento da linguagem e vários deles têm aplicado ferramentas

neurofisiológicas nessa investigação. Vamos discutir alguns desses estudos e tentar compreender melhor a relação entre prosódia e processamento de sentenças; mas, antes, devemos estabelecer quais aspectos prosódicos serão considerados em nosso próprio estudo e buscar compreendê-los melhor.

Mattoso Camara Jr. (2000, p. 69) observa que "a apresentação do vocábulo na escrita se faz pelo critério formal. Deixa-se entre eles, obrigatoriamente, um espaço em branco". Ele mostra que escrevemos "proscrever uma lei", apesar de falarmos "proskreverumalei", e chama esse conjunto de palavras de "grupo de força". Para atualizar este e outros conceitos, vamos recorrer à terminologia proposta por Barbosa (2019), que diz que a prosódia segmenta e organiza nossos enunciados em "unidades prosódicas" e essas unidades prosódicas distribuem-se em vários níveis, da sílaba ao enunciado.

Em contraste com a palavra gráfica — que Mattoso Camara Jr. (2000) chama de vocábulo —, existe a palavra fonológica, considerada como um conjunto de sílabas que contém uma única sílaba tônica e que forma uma unidade prosódica. Aproveitando o exemplo de Mattoso Camara Jr., poderíamos considerar o sintagma "uma lei" como uma palavra fonológica.

Um nível acima da palavra fonológica, encontra-se o grupo acentual, que reúne mais de uma palavra fonológica, mas onde apenas uma das sílabas recebe o acento primário, destacando-se em meio às outras. Voltando ao exemplo de Mattoso Camara Jr., o que ele chama "grupo de força" poderia ser considerado um grupo acentual, como em "Meu objetivo é proscREVERumalei", ou "Eu gostaria de proscREVERumaLEI", onde as maiúsculas indicam a sílaba tônica primária do grupo acentual. É importante considerar o comentário de Barbosa (2019) quando diz que certos autores não diferenciam palavra fonológica de grupo acentual, o que sugere que essas fronteiras conceituais não são absolutas.

A unidade seguinte, o sintagma entoacional, diferencia-se do grupo acentual por ser definido pela entoação, em vez de por critérios de duração e intensidade. A entoação é resultado da variação da frequência fundamental (F0), que, dito de outra forma, seria equivalente à variação melódica (mais agudo ou mais grave) da voz, em contraste com as variações rítmicas (duração e intensidade) que caracterizam o grupo acentual.

A unidade de hierarquia mais alta, por assim dizer, seria o enunciado (também chamado "enunciado entoacional"), um ato de fala completo. Trata-se de um conceito prosódico que nos remete a Izre'el, 2018, citado acima, e sua ideia de "domínio predicativo", que estamos chamando, neste trabalho, de predicação. Seu limite é definido por uma "fronteira terminal", cuja definição explícita

daremos adiante. O grupo de força de Mattoso Camara Jr. poderia, em situação discursiva apropriada, corresponder a um enunciado:

— O que somos?

— Anarquistas!

— E o que queremos?

— ProscreeverumaLEI!

Importante notar que as duas variáveis que escolhi chamar de ritmo e melodia são tratadas de forma mais técnica pelos estudiosos da prosódia. A prosódia é definida como tal a partir de suas funções linguísticas, ou seja, de acordo com as **percepções** que os falantes da língua têm dos fenômenos físicos que a caracterizam. Esses fenômenos físicos, a que Barbosa (2019) denomina correlatos físicos da prosódia são, principalmente, a frequência fundamental, a duração e a intensidade. A frequência fundamental (F0) é produzida pela vibração das pregas vocais e é medida em número de vibrações por segundo, representadas em Hertz (Hz), ou em semitons, como na música. A duração é medida pelo tempo de realização de determinados eventos fonológicos, como a sílaba ou os grupos acentuais. Um exemplo de uso da duração com efeito prosódico seria a emissão da palavra "madeeeeiira" com duração aumentada da sílaba "dei". O último correlato físico da prosódia que mencionaremos é a intensidade sonora, medida em decibéis (dB). Há diversos aspectos desses correlatos físicos da prosódia que não vamos mencionar aqui, pois nosso estudo não lidará diretamente com eles, mas com seus correspondentes perceptivos.

Os fenômenos sonoros físicos são percebidos pelo intérprete por intermédio do processamento dos sinais transmitidos pelo sistema auditivo para o sistema nervoso. O resultado é que não há correlação direta entre manifestação física e manifestação cognitiva dos elementos prosódicos, apesar de ser possível identificar certos correlatos perceptivos resultantes dos correlatos físicos mencionados acima. Esses correlatos perceptivos seriam o *pitch*, a duração percebida e o volume. O *pitch* é a percepção da F0 como um som mais grave ou mais agudo. A duração percebida é a percepção da duração do som como sendo mais curto ou mais longo. E o volume é a percepção da intensidade do som como sendo mais forte ou mais fraca (*loudness*). É importante frisar que a percepção humana, de um modo geral, não percebe cada um desses fenômenos — *pitch*, duração percebida e volume — de maneira absoluta, mas relativa aos fenômenos prosódicos que os circundam. Ou seja, um segmento

costuma ser percebido como mais ou menos duradouro, mais ou menos intenso, ou mais ou menos agudo ou grave se for mais ou menos duradouro, mais ou menos intenso, ou mais ou menos agudo ou grave em relação aos elementos adjacentes ou próximos com os quais se relaciona. Em outras palavras, os eventos prosódicos são percebidos e interpretados dentro de contextos prosódicos específicos.

Neste momento, poderíamos voltar a citar Platão (que talvez tenha sido o primeiro a empregar o termo grego "prosódia", significando "junto — ou em direção — ao canto"), retomando a analogia entre prosódia e música. Mas recorreremos a Plínio (BARBOSA, 2019), que leciona que a prosódia não considera diretamente o conteúdo segmentar da língua (as palavras e seus componentes fonológicos necessários, ou aquilo que é dito), mas o componente suprasegmentar (ou "como" as coisas são ditas). Ela teria duas funções primordiais: a marcação de proeminência — que dá relevo a determinado segmento do enunciado —; e a segmentação do próprio enunciado — que revela ao intérprete os limites do enunciado e de suas unidades prosódicas internas.

Barbosa e Raso (2018) nos lembram de que o único correlato gráfico da prosódia registrado na escrita é — talvez — a pontuação. E lembram, também, que a própria fala, se não for registrada de algum modo, se perde logo após a emissão. O que fica registrado na mente do intérprete é uma representação cognitiva que não equivale literalmente à fala que lhe deu origem. Hoje em dia, entretanto, é possível registrar a emissão dos enunciados da língua falada em suas dimensões sonora e visual, por meio de gravações de áudio e vídeo; e é possível, também, registrar e estudar, minuciosamente, os correlatos físicos de seus aspectos prosódicos.

O aspecto mais relevante da prosódia para este trabalho é a segmentação da fala lida. Por meio da prosódia, o emissor segmenta o fluxo — de outra forma, contínuo — da fala em unidades que são percebidas pelo interprete como distintas. Essa segmentação é muito importante para a compreensão da fala. Não há como negar a necessidade de sinalizar, por exemplo, o final de um enunciado e, conseqüentemente, a possibilidade de iniciar outro. Ou de, em muitas situações, segmentar um enunciado em unidades menores, indicando para o intérprete a organização interna desse enunciado. Ou mesmo a obrigatoriedade de não segmentar determinados trechos da fala, sob pena de mudar seu significado ou tornar o enunciado ininteligível (FRAZIER; CLIFTON; CARLSON, 2004).

As abordagens teóricas sobre a origem dessas segmentações, ainda segundo Barbosa e Raso (2018), é variada, podendo ser, por exemplo, sintática, informacional ou conversacional. O estudo dos parâmetros físicos objetivos (F0, duração e intensidade) associados aos locais onde os intérpretes

identificam, perceptualmente, as fronteiras entre segmentos, revela grande variedade de combinações e pesos, tanto entre intérpretes, como entre línguas ou circunstâncias distintas de emissão da fala. Independentemente dessa variedade, a percepção dos intérpretes a respeito dessas segmentações, especialmente ao final de enunciados, é bastante congruente. E, para os objetivos deste trabalho, essa congruência perceptual da segmentação da fala será muito mais relevante do que a variabilidade dos parâmetros físicos sonoros que lhes dão origem.

Diferentes teóricos defendem a existência de diferentes fronteiras prosódicas, com diferentes funções, mas parece ser possível considerar que haja, universalmente, pelo menos dois tipos principais de fronteiras: terminais e não terminais. Uma fronteira terminal indica que o trecho atual do discurso (ou o enunciado, ou a sentença) foi concluído. As fronteiras não terminais, por outro lado, indicam uma segmentação interna do trecho atual do discurso (uma espécie de constituinte prosódico da sentença), mas avisam, também, que aquele trecho (enunciado, sentença etc.) ainda não terminou. Ou seja, sinalizam para o intérprete algo como: "preste atenção no que eu disse até agora, mas fique comigo porque ainda não terminei".

As fronteiras terminais e não terminais são necessárias à comunicação e são produzidas com intenção comunicativa — apesar de, muitas vezes, involuntariamente — pelo emissor da fala. Há outros fenômenos, entretanto, que são acidentais, e interrompem o fluxo normal da fala sem intenção comunicativa explícita, como a interrupção (que chamaremos de pausa) e a hesitação. A pausa, como o nome já diz, acontece quando o emissor, por qualquer motivo, interrompe o fluxo da fala sem que a interrupção tenha relação com o discurso. Ocorre, por exemplo, quando o falante está elaborando o conteúdo da sequência de sua fala. A hesitação é uma espécie de interrupção preenchida por sons, geralmente vocálicos ("o livro...eeeeeh...do Plínio"), também chamada de pausa preenchida.

O'Connell e Kowal (2008) defendem que as hesitações e outros fenômenos correlatos normalmente considerados como disfluências são, na verdade, parte integrante da fluência comunicativa da fala espontânea. A linguística tradicional, segundo eles, endossaria a visão chomskiana de "erros de performance" e isso seria consequência de uma ilusão teórica derivada do conceito de competência. Para eles, a fluência genuína incluiria fenômenos como hesitações, pausas, preenchimentos, repetições, falsos inícios, interrupções e falas simultâneas como instrumentos genuínos para sua realização. Um exemplo da impossibilidade fisiológica de uma fala contínua, como pretendem certos teóricos, seria a própria respiração, que interrompe o fluxo da fala a cada cinco segundos, em média, de acordo com os autores.

Frazier, Clifton e Carlson (2004) dizem que há margem para muita opcionalidade na segmentação prosódica dos enunciados, com várias possibilidades alternativas não prejudiciais ao significado ou à compreensão das sentenças. O que importa, segundo eles, não é a presença ou ausência de uma determinada fronteira prosódica, mas a relação harmônica entre os diversos fenômenos prosódicos presentes em um determinado enunciado. Eles testaram algumas hipóteses sobre opções em escolhas prosódicas e detectaram a importância de, pelo menos, dois fatores nessas escolhas. A primeira hipótese testada confirmou que os intérpretes aceitam fronteiras não terminais antes de constituintes prosódicos longos. A segunda é que os intérpretes não aceitam constituintes prosódicos semanticamente incoerentes. Por exemplo, nas sentenças abaixo, (23) e (24) seriam aceitáveis, mas (25), não:

(23) A lady from the accounting office visited...¹

(24) A lady / from the accounting office / visited...

(25) A lady / from the accounting office visited...

As sentenças acima mostram que é possível não inserir fronteiras não terminais, como em (23), ou inseri-las abundantemente, como em (24). O importante seria a coerência semântica dessas inserções. Uma inserção como em (25), segmentando o argumento e não acrescentando nenhuma outra fronteira na sentença, indicaria que a fronteira inserida (a única) é a mais relevante da sentença e que essa quebra do argumento deve ser importante. No entanto, não é possível para o intérprete compreender a motivação dessa fronteira. Em (24), por outro lado, como há outra fronteira separando o argumento do verbo, a fronteira inicial, que segmentou o argumento, perderia sua importância relativa e o conjunto de fronteiras passaria a ser harmônico.

¹ Uma mulher da contabilidade visitou... (tradução nossa)

2.10 Retomando alguns conceitos — prosódia, constituintes sintáticos, estrutura argumental, *chunking* — e vendo se é possível empacotá-los em outro: o *closure positive shift*

Friederici (2017) faz uma boa revisão da prosódia sob o ponto de vista dos estudos neurofisiológicos. Ela diz que as fronteiras prosódicas são extremamente relevantes na compreensão de "quem está fazendo o que para quem". Em outras palavras, elas são importantes nas decisões tomadas pelo intérprete sobre a estrutura argumental da sentença.

Friederici e outros autores identificaram, em 1999, por meio de ERP, uma onda positiva centro-parietal que se correlacionava com a presença de uma fronteira prosódica não terminal. Eles deram a essa onda o nome de *closure positive shift* (CPS), porque a onda se associava ao encerramento de um constituinte prosódico e os autores acreditavam que ela fosse um marcador do processamento da fronteira. Estudos posteriores demonstraram que, além do alemão, língua do estudo original, o CPS estava presente em diversas outras línguas. Demonstrou-se, também, que o CPS é sensível a qualquer tipo de marcação de fronteira: pausa, modulação de duração e variação de F0. Além disso, ela está presente em sentenças deslexicalizadas, nas quais, por meio de processamento do áudio original, anula-se o conteúdo segmentar (as palavras), mantendo-se apenas o suprasegmentar (os aspectos prosódicos da sentença). E, curiosamente, o CPS também surge durante a leitura, na posição de vírgulas que marcam fronteiras não terminais. A interpretação atual do fenômeno, segundo Friederici (2017), é que a CPS assinala o fechamento, o encerramento, a conclusão ("*closure*", no original) do constituinte prosódico cujo término é marcado pela fronteira respectiva. Ele não seria uma resposta às manifestações acústicas da prosódia, mas ao seu significado linguístico: a marcação/delimitação de um constituinte.

Diversos estudos demonstram que a presença do CPS se correlaciona com fenômenos aparentemente distintos. Männel e Friederici (2016) estudaram crianças alemãs de três anos de idade e, ao contrário de adultos e de crianças a partir de seis anos de idade, nos quais o CPS é desencadeado por qualquer pista prosódica marcadoras de fronteira, as crianças de três anos apresentam CPS em resposta a fronteiras marcadas com pausa e aumento da duração silábica pré-fronteira, mas não àquelas marcadas apenas com entoação. As autoras levantam algumas hipóteses: deve haver diferenças entre as línguas com relação às pistas prosódicas a que as crianças reagem inicialmente, ao longo dos primeiros anos de vida; e, no caso das crianças alemãs, a insensibilidade inicial à entoação dever ser

consequência da importância relativa da entoação na segmentação da língua alemã (no alemão, ao contrário do inglês, a entoação seria menos importante do que aspectos rítmicos).

Com relação ao PB, Musiliyu et al. (2019) demonstraram a evocação do CPS em fronteiras terminais e não terminais em sentenças faladas como "Assim que Paula viu sua amiga / ela fechou a janela. // Foi abrir a porta." (onde "/" indica fronteira não terminal e "//" indica fronteira terminal). Um achado interessante de seu estudo foi que as fronteiras não terminais evocaram reações eletrofisiológicas mais precoces, mais duradouras e de maior amplitude do que as fronteiras terminais.

Bögels et al. (2011), em revisão sobre a importância das fronteiras prosódicas no processamento de sentenças avaliadas por ERP, lembram que as fronteiras prosódicas frequentemente coincidem com as fronteiras sintáticas, mas que a correlação não é absoluta. Além disso, o CPS teria sido evidenciado em associação com a leitura de constituintes muito longos em coreano, mesmo sem a presença de vírgulas, e também em frases musicais. A conclusão dos autores é que o CPS refletiria algum tipo de estruturação ou fraseamento genérico do *input* de informações, e não seria um fenômeno exclusivamente linguístico.

Hilton et al. (2019) prestaram atenção aos relatos de evocação de CPS em situações menos ortodoxas, como frases deslexicalizadas, leitura silenciosa, sentenças formadas por pseudopalavras e frases musicais; e resolveram investigar a hipótese, levantada por outros, de que o CPS poderia refletir um domínio cognitivo geral, não necessariamente linguístico, como se fosse a delimitação de uma estrutura perceptual qualquer agrupada ("*closure of a grouped perceptual structure*", no original). Eles analisaram o ERP em participantes expostos a dois experimentos: a) sequências de três verbos (escolhidos entre girar, rolar, balançar, puxar e pegar), com ou sem fronteira prosódica entre o segundo e o terceiro verbo (ex: girar, rolar...., e balançar); e b) sequências de três movimentos em que um ator executava as mesmas ações descritas pelos verbos, manipulando com as mãos um pequeno objeto, com ou sem uma pausa entre o segundo e o terceiro movimentos. O estudo demonstrou que em ambos os casos — linguístico e cinemático (visual) — a fronteira (prosódica ou cinemática) evocou um CPS muito semelhante.

Retomando o modelo "*multiscale information transfer framework* (MSIT)", de Blumenthal-Dramé e Malaia (2018), discutido anteriormente, Hilton et al. (2019) também levantam a hipótese de que o CPS refletiria um processamento de fronteiras de domínio cognitivo geral e de que os mecanismos de segmentação de sinais linguísticos e motores poderiam ser os mesmos.

Em suma, parece haver evidências de que os constituintes (argumentos, *grammatical relations*, agente, pacientes, enfim, os componentes da **predicação**) de uma sentença, delimitados (ou não) no fluxo da fala por elementos prosódicos, e cognitivamente aglomerados (*chunking*) para formar unidades de informação e liberar espaço na WM, podem ser, enfim, considerados como realidades cognitivas, pois parecem ter uma assinatura neurofisiológica detectável no tecido cerebral: o CPS.

2.11 O corpus e as incongruências: fala lida, fala espontânea, fala lida espontânea...

Abordando agora aspectos relacionados ao corpus que será objeto deste estudo — um corpus de leituras em voz alta — Merlo (2012) faz algumas considerações sobre as diferenças entre fala lida e fala espontânea e levanta argumentos que vão ao encontro de O'Connell e Kowal (2008), já discutidos nessa revisão. Ela faz críticas à ideia comum de que "o correto é falar como se escreve". Em seu estudo, pausas (do tipo “demarcativas”) e hesitações revelaram-se perfeitamente compatíveis com (e auxiliares de) a fluência na fala espontânea. A crítica é relevante para este estudo por que lidamos com questão correlata, que passa pela definição do que seria adequado, congruente, ou inadequado, incongruente, em falas lidas. Dessa forma, consideramos que não há uma forma padrão única, ideal para a leitura de uma sentença, mas um conjunto de possibilidades, igualmente aceitáveis, atendendo um critério maior que é a fidelidade a seu conteúdo semântico.

De acordo com proposta de classificação de corpora elaborado por Barbosa (2012), em que se distribuem as amostras em eixos ortogonais segundo "grau de controle do experimentador" e "gênero/tipo de corpus", os pronunciamentos parlamentares lidos (nosso corpus) poderiam ser classificados como "leitura em voz alta espontânea". Eles seriam classificados como um tipo de fala espontânea porque, apesar de serem baseadas em leitura de textos, a seleção dos textos não sofreu nenhuma interferência do experimentador, ou de quem analisa o corpus.

Os textos deste corpus são escritos com o objetivo de serem lidos. Portanto, quem os escreveu teve uma preocupação com a oralidade (escrever como se fala) e com a compreensibilidade para quem ouve a leitura, o que não se observa em textos escritos para uma leitura silenciosa, como esclarece Lehrman (2010). Do ponto de vista da análise de incongruências neste corpus, pretendemos seguir a linha de Merlo (2012) e O'Connell e Kowal (2008), utilizando critérios de exclusão de incongruências bastante amplos e identificando trechos com incongruências apenas quando não puderem ser classificados como normais, nem na fala lida, nem na fala espontânea. Em suma, pretendemos

considerar incongruentes apenas os trechos que não se encaixarem em: a) o que se espera da fala lida; ou b) o que se considera normal na fala espontânea.

2.11.1 Especificidades do processo de leitura e suas consequências para a fala lida

Kondo e Mazuka (1996) levantam outro ponto relevante com respeito à leitura, que é a produção da prosódia na leitura em voz alta. Eles relatam estudos que mediram o que se chama de "*eye-voice span*" (a distância em caracteres entre o foco da visão no texto lido e o que está sendo falado em determinado momento) em japonês e revelaram que essa distância é de apenas 2,5 caracteres. Fazendo uma extrapolação dos caracteres da escrita em japonês para a escrita em inglês, eles calculam uma relação de 1 para 2, ou seja, 2,5 caracteres japoneses corresponderiam a cerca de 5 letras em inglês. A conclusão dos autores é que a visão dos leitores está apenas alguns caracteres à frente do que está sendo falado, sugerindo que a prosódia produzida não se baseia em uma visão global da sentença, mas apenas em pistas locais, resultando em eventuais erros e reparos.

Confirmando essa impossibilidade de se ter uma visão global do período durante a leitura (a visão é local), Dehaene (2009) pontua que nossos olhos não conseguem enxergar mais do que 10 a 12 letras em cada sacada (a cada fixação), algo entre 3 e 4 à esquerda, e de 7 a 8 à direita do ponto de fixação. Essa proporção se aplicaria a sistemas de escrita como o nosso, lidos da esquerda para a direita, e se inverteria em sistemas onde se lê da direita para a esquerda. Em sistemas de escrita com maior densidade de informação por caractere, como o chinês e o japonês, as sacadas seriam mais curtas e o número de letras percebidas a cada fixação seria proporcionalmente menor, como mostrou o estudo japonês acima.

Duas questões adicionais relativas ao leitor que merecem ser consideradas podem ser descritas como: a) não enxergar o que está escrito; e b) enxergar o que não está escrito. São questões relativas ao processamento cognitivo, não ao sistema visual. O leitor — quem lê o texto em voz alta — é, ao mesmo tempo, intérprete e emissor da fala. O resultado de todo o processo de interpretação e emissão transparece para a audiência sob a forma de fala lida. Na etapa de interpretação por parte do leitor, ele está sujeito aos dois fenômenos citados acima. O primeiro item (não enxergar o que está escrito) foi abordado por Christiansen e Chater (2015) e se chama "*innatentional blindness*". É resultado direto do processo de compressão e representação seletiva do fluxo de informações obtidas durante o processamento das sentenças. Esse fenômeno é de domínio geral, não sendo privativo da linguagem. Um exemplo dramático foi dado por Simons e Levin (1998) em um experimento onde um pesquisador

pedia informações na rua para uma pessoa qualquer. Enquanto a pessoa dava as informações, o pesquisador era substituído por outro sem que a pessoa percebesse que a troca estava acontecendo. O surpreendente é que a pessoa continuava dando as informações para o novo pesquisador sem “enxergar” que se tratava de outro indivíduo.

No fenômeno seguinte — "enxergar" o que não está escrito —, os leitores, muitas vezes, trocam uma palavra que está escrita no texto por outra que não está escrita. De acordo com a literatura, pode haver, pelo menos, duas razões para isso. A primeira comunica-se com o conceito de "*entrenchment*", já abordado neste texto. Langacker (2017, p. 41) observa que "*at some point, a unit is well enough established (deeply enough entrenched) that it constitutes 'an event waiting to happen': once initiated, it will run to completion if not deflected from this course*"¹. Ou seja, a palavra que foi pronunciada, mas que não estava escrita, estaria mais *deeply enough entrenched* ² no contexto em que foi pronunciada.

A outra ideia tem a ver com o conceito de predição ("*prediction*"), também já abordado nesta revisão. O que merece ser acrescentado é a possível relação entre predição e o modelo proposto por Pickering e Garrod (2013). Esses autores, como já discutimos, propõem que o sistema de processamento de sentenças depende do sistema de produção. Esses dois sistemas, portanto, funcionariam em conjunto, como se fossem apenas um. O intérprete se socorre de seu próprio sistema de produção da língua para predizer o restante da sentença que está interpretando (ouvindo ou lendo). Extrapolando esse raciocínio para o caso da leitura em voz alta, onde o intérprete recorre duas vezes ao sistema de produção — a primeira para predizer, e a segunda para produzir —, predições incorretas fariam com que ele "enxergasse" no texto algo que não está lá na realidade.

¹ A partir de certo momento, uma unidade está tão bem estabelecida (tão profundamente entranhada), que ela constitui um ‘evento prestes a acontecer’: uma vez iniciada, e não sendo desviada de seu curso, ela ‘acontecerá’ por completo. (tradução nossa)

² Suficientemente entranhada. (tradução nossa)

3. METODOLOGIA

Esta pesquisa foi submetida a avaliação pelo Comitê de Ética em Pesquisa da Faculdade de Ceilândia da Universidade de Brasília, sendo registrada sob o número CAAE 17295019.3.0000.8093. O parecer consubstanciado n° 3.526.538 aprovou o protocolo de pesquisa na data de 22 de agosto de 2019, em consonância com a Resolução 466/12 do Conselho Nacional de Saúde.

O objetivo geral da pesquisa é descobrir se há correlação entre a) incongruências eventualmente observadas na leitura de textos e b) aspectos linguísticos específicos destes textos. Nossa hipótese de trabalho é a de que tal correlação deve existir. Para atender a esse objetivo e testar a hipótese, buscamos o desenho estatístico mais adequado para o estudo. Se fizéssemos um estudo experimental, cujas variáveis pudéssemos controlar, teríamos duas categorias de variáveis: dependentes (as incongruências) e independentes (aspectos linguísticos dos textos). Entretanto, não se trata de um estudo experimental.

O que faremos, então, será identificar as incongruências e, a seguir, identificar o ambiente linguístico dos textos onde elas ocorreram ou não ocorreram. Trata-se, portanto, de identificar os **casos** (incongruências) e, posteriormente, seu o ambiente linguístico. E, para efeitos estatísticos, será necessário identificar **controles** (situações em que não houve incongruências) que sejam comparáveis (o mais amplamente possível) aos casos. Trata-se, portanto, de um estudo que se assemelha aos estudos do tipo caso-controle.

Nesse tipo de estudo, uma escolha adequada dos controles é fundamental. Acreditamos que os controles mais comparáveis são produzidos pelos próprios indivíduos que geraram os casos. Um desenho para esse tipo de estudo foi proposto por Maclure (1991) e revisado por Maclure e Mittleman (2000), e é denominado “caso-crossover”.

Num estudo caso-crossover, como os indivíduos são os mesmos tanto na situação de caso como de controle, considera-se, por analogia com estudos experimentais, que os indivíduos são controles de si mesmos e que houve um crossover com relação à exposição. O que é feito, então, é avaliar o desfecho (o evento pesquisado, que, em nosso caso, são as incongruências) em duas situações: indivíduo A exposto *versus* indivíduo A não exposto.

Em nosso estudo, o mesmo indivíduo, no mesmo pronunciamento, foi responsável por gerar eventos (incongruências) e não eventos (leitura congruente). Nesses dois tipos de situação (eventos e

não eventos), foi analisado o grau de exposição a determinados elementos linguísticos, selecionados, dentre os citados na literatura revisada, como tendo maior potencial para induzir incongruências. Infelizmente, nem sempre teremos um número suficiente de eventos para que o próprio indivíduo seja, ao mesmo tempo, caso e controle. Nessas situações, a análise será feita agrupando os eventos produzidos por todos os participantes.

Deve-se notar que determinadas incongruências entre a fala lida e o texto escrito podem ser intencionais, voluntárias, revelando escolhas conscientes dos falantes. Nesses casos, acreditamos que tais incongruências não trarão prejuízo à fala lida. Por outro lado, quando a incongruência for prejudicial à fala lida, acreditamos que ela foi produzida de forma involuntária. Portanto, em nossas análises, procuraremos distinguir as incongruências que não trazem prejuízo à fala lida (a leitura mantém sua adesão ao sentido do texto) e excluí-las de determinadas análises. E vale a pena esclarecer que consideramos “prejuízo” tudo aquilo que afasta a fala lida da intenção comunicativa original, do sentido evidente do texto escrito.

A seguir, descrevem-se os detalhes metodológicos do trabalho e abordam-se, em sequência, cada um dos tópicos abaixo:

- 3.1 - O corpus;
- 3.2 - Os indivíduos (os falantes) que geraram o corpus;
- 3.3 - Os critérios de seleção para extração de pronunciamentos do corpus;
- 3.4 - As características do corpus e dos falantes após a seleção (ou seja, os pronunciamentos que foram efetivamente analisados e os indivíduos que os produziram);
- 3.5 - Os métodos utilizados para identificação das incongruências
- 3.6 - Os critérios e as categorias utilizadas para classificar as incongruências observadas;
- 3.7 - O método de validação dos critérios e categorias utilizadas para identificação das incongruências;
- 3.8 - Os tipos de análise realizados para caracterizar os ambientes linguísticos (exposição) em que as incongruências ocorreram e não ocorreram (eventos e não eventos); e
- 3.9 - Aspectos metodológicos complementares

3.1 O corpus

A fonte de dados deste estudo é um corpus de domínio público (BRASIL, Senado Multimídia, 2019) composto por pronunciamentos parlamentares proferidos no Plenário do Senado Federal. Os pronunciamentos são monólogos de caráter político sobre assuntos diversos, eventualmente interrompidos por apertes. Podem incluir trechos de fala lida e/ou fala espontânea e são registrados da seguinte forma:

1. São gravados em áudio e vídeo por equipe técnica da instituição. Os arquivos estão disponíveis para download em formato MP3 (áudio) com compressão média de 1 Mb/minuto; e em formato MPEG-4 (vídeo) com compressão média de 3,7 Mb/minuto (BRASIL, Senado Multimídia, 2019).
2. São transcritos por taquígrafos profissionais da instituição durante a produção das falas, muitas vezes com auxílio dos textos escritos em que se basearam. O protocolo para a produção das notas taquigráficas prevê: 1) a transcrição deve reproduzir o mais fielmente possível o que foi dito; e 2) eventuais equívocos involuntários produzidos pelos falantes (ex: desvios de ortoépia e prosódia) são, via de regra, corrigidos. Os textos das transcrições taquigráficas também estão disponíveis para download (BRASIL, Atividade legislativa: pronunciamentos, 2019).
3. Os textos escritos que servem de base aos pronunciamentos ficam armazenados em sistema informatizado interno da instituição, sob a responsabilidade de órgão que presta serviços de apoio aos pronunciamentos (quando tais serviços são solicitados).

3.2 Os falantes

São 39 cidadãos brasileiros, de ambos os sexos (32 homens, 7 mulheres), com 35 anos de idade ou mais (média de 63 anos), com nível de escolaridade predominantemente superior, que exerceram mandato parlamentar e proferiram discursos lidos no Plenário da instituição nos anos de 2017 e 2018. Os estados natais dos participantes representam a variedade do PB: 4 de Goiás, 4 de São Paulo, 3 da Paraíba, 3 de Santa Catarina, 2 de Alagoas, 2 do Ceará, 2 do Mato Grosso do Sul, 2 de Minas Gerais, 2 do Pará, 2 do Rio de Janeiro, 2 do Rio Grande do Norte, 2 do Rio Grande do Sul, 2 de Roraima, e 1 do Acre, Bahia, Maranhão, Mato Grosso, Pernambuco, Piauí e Sergipe.

3.3 Os critérios de seleção dos pronunciamentos analisados

Foram selecionados todos os pronunciamentos proferidos nos anos de 2017 e 2018 para os quais havia texto escrito disponível no sistema interno da instituição. O processo de extração dos pronunciamentos dos bancos de dados seguiu os seguintes passos:

1. Foram baixados todos os textos escritos relativos aos anos de 2017 e 2018 que estavam disponíveis no banco de dados interno da instituição.
2. Foi utilizado o programa Octoparse (OCTOPARSE, 2019) para extrair as URLs e baixar do site da instituição todos os arquivos de transcrições taquigráficas relativos aos anos de 2017 e 2018.
3. Foi utilizado o programa Plagiarism Checker X (PCX, 2019) para fazer o cruzamento de informações entre os textos escritos e as transcrições taquigráficas e detectar coincidências, indicando os pronunciamentos em que houve leitura dos textos escritos ou de parte deles (se qualquer parte do texto escrito é lido no Plenário da instituição, a transcrição taquigráfica registra esse evento).
4. Foram baixados os arquivos (vídeo, áudio, texto escrito e apanhamento taquigráfico) de todos os pronunciamentos em que houve leitura dos textos escritos.

3.4 Características do corpus e dos falantes após aplicação dos critérios de seleção

Aplicados sucessivamente os critérios de seleção descritos acima ao corpus original, obtivemos o seguinte corpus pós-seleção:

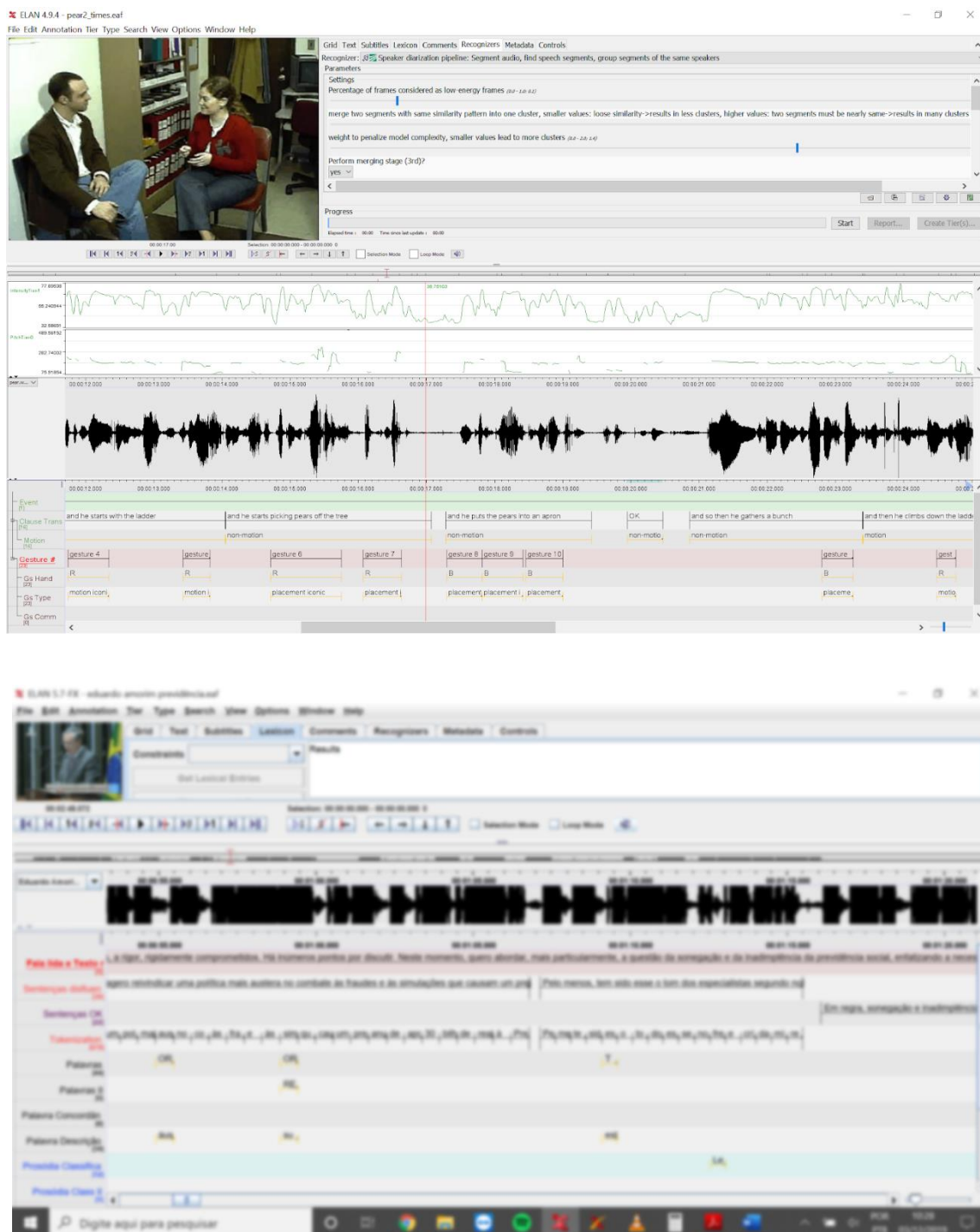
- a. 1610 pronunciamentos (textos escritos)
- b. 280 pronunciamentos (dos 1610 textos extraídos) com trechos lidos
- c. 39 falantes distintos que leram ao menos um trecho de um pronunciamento
- d. Para cada falante, selecionamos, dentre os 280 pronunciamentos com leituras, o que tinha a quantidade de material lido (um texto por falante).

- e. Os 39 pronunciamentos selecionados corresponderam a 8 horas e 28 minutos de áudio/vídeo.
- f. Nessas 8 horas e 28 minutos de áudio/vídeo selecionados, identificamos 5 horas de fala lida.
- g. Essas 5 horas de fala lida foram segmentadas em períodos com e sem incongruências.
- h. Estes e os demais resultados do trabalho estão descritos e analisados com mais detalhe no capítulo Resultados.

3.5 A identificação das incongruências

Os arquivos de áudio, vídeo, e textos de cada pronunciamento selecionado foram analisados no programa ELAN (MAX PLANCK INSTITUTE FOR PSYCHOLINGUISTICS, 2019), que permite fazer anotações ilimitadas, hierarquizadas e sincronizadas com qualquer um dos arquivos de cada pronunciamento. O programa foi instalado em computadores do tipo PC, com sistema operacional Windows 10 Home de 64 bits. A interface do programa é reproduzida, a título de ilustração, na **Figura 1**, onde há duas imagens: a primeira foi retirada do site do programa; e a segunda foi extraída durante as análises para este trabalho e propositalmente turvada em software de processamento de imagens para não revelar a identidade do emissor do pronunciamento.

Figura 1— Aparência em tela do software de anotação de dados ELAN



O processo de detecção, classificação e registro de incongruências consistiu em ler o texto escrito e, na sequência, assistir ao respectivo trecho do pronunciamento (fala lida) com o vídeo em janela maximizada e o áudio em headphone Bose QuietComfort®20. Cada incongruência identificada foi anotada no ELAN e classificada de acordo com as categorias descritas na próxima seção.

Adicionalmente, anotamos também os eventos respiratórios produzidos pelos falantes. Por meio de uma observação conjunta do áudio (som de inspiração/expiração) e do vídeo (movimentos torácicos, bucais, nasais e cefálicos), identificamos e anotamos no ELAN os momentos em que os falantes respiraram. Com base nessas anotações, obtivemos duas variáveis:

- Tempo entre pausas respiratórias (em segundos): total de pausas respiratórias dividido pelo tempo total de fala de cada falante (note-se que se trata do tempo transcorrido entre o início de dois movimentos respiratórios sucessivos, incluindo a pausa respiratória, que seria correspondente ao tempo de inspiração)
- Palavras entre pausas respiratórias: número de palavras produzidas por cada falante durante a leitura dividido pelo total de pausas respiratórias

Para efeito de análise, correlacionamos essas variáveis das seguintes formas: entre si; com duas variáveis “independentes”, idade e sexo; e com a variável taxa de elocução, descrita adiante.

3.6 As incongruências

Cada incongruência observada foi anotada no respectivo *tier* do ELAN e distribuída em categorias previamente especificadas. A escolha das categorias baseou-se em uma abordagem mista: empírica e teórica. A face empírica contempla incongruências observadas de forma assistemática e casual no dia a dia do trabalho do pesquisador; e, de forma mais sistematizada, em uma fase preparatória para o início deste estudo. A face teórica volta-se para certos desvios costumeiramente mencionados na literatura linguística. Entretanto, não aderimos à classificação de disfluências tipicamente encontrada na literatura fonoaudiológica, como a proposta por Schiefer e Arcuri (2014), pois acreditamos que nossa abordagem está buscando fenômenos diversos. Nós lidamos exclusivamente com incongruências entre texto escrito e leitura em voz alta. As incongruências que procuramos identificar — sejam segmentares (o que foi dito diverge do que estava escrito), sejam suprasegmentares (o que foi dito diverge do que o pesquisador considera prosodicamente adequado à reprodução do trecho escrito) — derivam das divergências observadas durante a leitura. A classificação descrita por Schiefer e Arcuri (2014), por outro lado, espelha o chamado ABFW, um teste padronizado de linguagem infantil cujo objetivo é identificar e classificar disfluências *stricto sensu*, ou seja, perturbações do fluxo da fala. Tal classificação separa as disfluências em típicas (comuns) e atípicas

(gagas) e, além de incluir disfluências que não são relevantes para nosso estudo, como as disfluências gagas (não estamos em busca de idiosincrasias de fala, mas de dificuldades universais de leitura), exclui incongruências que podem não alterar o fluxo normal da fala, mas que afastam a fala lida do texto escrito. A existência de um texto escrito como gabarito permite identificar incongruências que não são, necessariamente, disfluentes, como é o caso das trocas de palavras, dos desvios de ortoépia e concordância, e das omissões de palavras (a palavra falada, ou não falada, diverge do texto escrito e pode, ou não, prejudicar o fluxo da fala e a mensagem pretendida); e das incongruências suprasegmentares, como inserções e deleções de fronteiras prosódicas e outras incongruências rítmicas e tonais (o “ritmo” e/ou a “melodia” da fala lida divergem do que o pesquisador considera minimamente adequado à expressão oral da mensagem escrita, mas podem não prejudicar o fluxo da fala; e, além disso, podem, eventualmente, passar despercebidos a quem não conhece o texto escrito).

A seguir, descrevemos as classificações de incongruências que empregamos neste estudo e as ilustramos com alguns fragmentos de áudios retirados do corpus:

Incongruências segmentares:

- **Troca de palavra:** a palavra falada é uma palavra do léxico, mas diverge da palavra escrita
 - Texto escrito: “...encaminhado à sanção..”. Fala lida: “...encaminhado à sessão...à sanção...”
 - Obs: neste caso, houve retratação, e a retratação incluiu a contração artigo+preposição (que não havia sofrido troca)
 - Áudio:
https://drive.google.com/file/d/1Phlv8X220dswdd_ZsLxDiX6HiKrhR4Em/view?usp=sharing
 - Texto escrito: “...para consumir o que necessitam...”. Fala lida: “...para conseguir o que necessitam...”
 - Áudio:
<https://drive.google.com/file/d/1ZO3W5tS9YWQC5E38PiPrvbbxkXPbFrGs/view?usp=sharing>
 - Texto escrito: “...é devolver...”. Fala lida: “...é desenvolver...devolver”
 - Obs: houve retratação da troca de palavra

- **Áudio:**
https://drive.google.com/file/d/1SOwqWdULmpfudxI_QY6OkBpK_ezn_w1SM/view?usp=sharing
 - Texto escrito: “...novas oportunidades...”. Fala lida: “...nossas...novas oportunidades...”
 - Obs: houve retratação da troca de palavra
 - **Áudio:** <https://drive.google.com/file/d/1X-6SBqB0Jccsb9um1AdUsCEotLj-tPoZ/view?usp=sharing>
 - Texto escrito: “...como grande parte...”. Fala lida: “...uma grande...como grande...”
 - Obs: houve retratação da troca de palavra
 - **Áudio:**
https://drive.google.com/file/d/1qzemoP6imvm3M7a9vuG2GWaF_VnDfYqnY/view?usp=sharing
- **Desvio de ortoépia¹:** a palavra falada não é uma palavra do léxico
 - Texto escrito: “...a corrupção e a economia...”
 - Obs: houve retratação do desvio de ortoépia
 - **Áudio:** <https://drive.google.com/file/d/1fYDeMXpGuK0QR7T-oskgwpfNbiwY8U3j/view?usp=sharing>
 - Texto escrito: “...chamou atenção, entretanto, ...”
 - Obs: houve retratação
 - **Áudio:** https://drive.google.com/file/d/1-SHBJrGMCD8vVIU43_YdtLfkU7bzoIOM/view?usp=sharing

¹ O termo “ortoépia” foi escolhido (entre outros possíveis) para indicar que o leitor errou a pronúncia da palavra original e a palavra resultante não pode ser considerada uma palavra do léxico do PB. Quando o erro do leitor resultou na pronúncia de uma palavra que está no léxico, o evento foi considerado como ‘troca de palavra’.

- **Omissão ou adição de palavra:** omitiu-se ou acrescentou-se uma palavra à fala, à revelia do texto escrito
 - **Omissão:** texto escrito: “...para que as crianças tenham...”. Fala lida: “...para as crianças...para que as crianças tenham...”
 - Obs: houve retratação com adição de palavra não omitida (“para”)
 - Áudio: <https://drive.google.com/file/d/1VikkYCXRlog0evwBdUZ8SFhCB1JUKoBO/view?usp=sharing>
 - **Adição:** texto escrito: “...são a ganância: a busca do lucro...; e a miséria: a busca da subsistência...”; Fala lida: “...são a ganância e a busca do lucro...”
 - Áudio: <https://drive.google.com/file/d/1tdL04TC6RYEJgBasOG-JYGPvz6Aan6vu/view?usp=sharing>
- **Desvio de concordância** (gênero e número): a palavra falada discorda em gênero ou número de outra(s) palavra(s) do período, e o texto escrito não continha tal discordância
 - **Singularização:** texto escrito: “...estão atolados...”. Fala lida: “...estão atolado...”
 - Áudio: https://drive.google.com/file/d/1WKJxFCdp3afwOPJm4u9Yfhq4gc_zwtjN/view?usp=sharing
 - **Singularização:** texto escrito: “...estudos divulgados, que já alertavam...”. Fala lida: “...estudos divulgados, que já alertava...”
 - Áudio: https://drive.google.com/file/d/1WHH8NoO2nsJhA72tK_JFKhLURrpzKieL/view?usp=sharing
 - **Masculinização:** texto escrito: “...reforma da previdência, com o destino da qual estamos...”. Fala lida: “...reforma da previdência, como o destino do qual estamos...”
 - Áudio: https://drive.google.com/file/d/1TNZEy7x7GWO_rTnZyZN3GTNCn8t1UZVn/view?usp=sharing
- **Incongruências segmentares não classificadas em outra parte:** anotação descritiva

- **Retratação:** releitura de trecho para correção de alguma(s) das incongruências acima
 - Obs: tal categoria é chamada, por Schiefer e Arcuri (2014), de “revisão”, e é caracterizada como “modificação no conteúdo da mensagem, na forma gramatical ou na pronúncia de uma palavra”. Aqui, entretanto, preferimos chamar o fenômeno de “retratação”, pois trata-se de uma revisão com o objetivo específico de corrigir uma determinada incongruência e retomar a aderência ao texto escrito, e não de revisar, de maneira genérica e inespecífica, o que foi dito.

Incongruências suprasegmentares:

- **Inserção (/) ou deleção (*) de fronteiras prosódicas:**
 - **Fronteira não terminal**
 - **Inserção:** o leitor excluiu palavras de um constituinte
 - **Deleção:** o leitor acrescentou palavra(s) a um constituinte
 - Exemplos de ambas as categorias:
 - Texto escrito: “Cerca de 75% da água do nosso País está localizada nos rios da Bacia Amazônica...”. Fala lida: “...rios/da bacia..”
 - Áudio:

https://drive.google.com/file/d/1RTaW6Oc74iBX_OapUBJBAUDVI6YuiRJB/view?usp=sharing
 - Texto escrito: “Na opinião de alguns, estados economicamente mais fracos ocupam espaços...”. Fala lida: “Na opinião de alguns * estados economicamente mais fracos / ocupam espaços...”
 - Áudio:

<https://drive.google.com/file/d/1DaW9QLuG-m9EkZ3WrqCFhL7vC6aqIWsZ/view?usp=sharing>
 - Texto escrito: “Contudo, é preciso atenção para os recalitrantes entraves...”. Fala lida: “Contudo * é preciso atenção / para...”

- **Áudio:** <https://drive.google.com/file/d/1TH5AMi-kVKGdIR1wjVTYOoS77k6388z2/view?usp=sharing>
- **Fronteira terminal**
 - **Inserção:** o leitor encerrou um período precocemente
 - **Deleção:** o leitor prolongou um período inapropriadamente
 - Exemplos de ambas as categorias:
 - Texto escrito: “Hoje, mais de 30 milhões de brasileiros estão atolados em dívidas impagáveis, segundo o Idec.”. Fala lida: “Hoje, mais de 30 milhões de brasileiros estão atolados em dívidas impagáveis/, segundo o Idec*.”
 - **Áudio:**
 - <https://drive.google.com/file/d/1dwTxByMFrIFLtyyamZxKxftpNOnZDRqSA/view?usp=sharing>
- **Pausa** (interrupção silenciosa e temporária do fluxo da fala) ou **hesitação** (pausa preenchida por som)
 - Pausa: texto escrito: “...faz mudar...”. Fala lida: “...faz... mudar...”
 - **Áudio:**
 - <https://drive.google.com/file/d/15SC2WEhcvkvgqmfrlTArBhlnHEoyOWb2/view?usp=sharing>
 - Pausa: texto escrito: “...atividades de formação continuada...”. Fala lida: “...atividades de formação... continuada...”
 - **Áudio:**
 - https://drive.google.com/file/d/1t9fvZA4rE1hJDOfr0dsk7zsexpZp5_qn/view?usp=sharing
 - Hesitação: texto escrito: “...juntas, as fábricas de fertilizantes...”. Fala lida: “...fábricas deee fertilizantes...”
 - **Áudio:**
 - <https://drive.google.com/file/d/1aw0tWcFaC3JMw4R6bBe5nqzyuuXL3iOY/view?usp=sharing>
 - Hesitação: texto escrito: “...comenta que estabilizamos...”. Fala lida: “...comenta queeeee estabilizamos...”.

- **Áudio:**

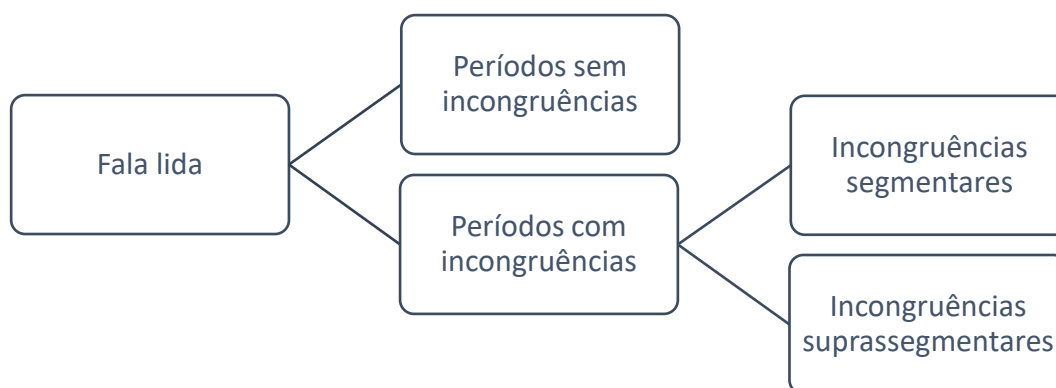
<https://drive.google.com/file/d/1BfBVkS0DaqfNzkFtHrqmgeaO4ZvtTwAM/view?usp=sharing>

- **Sentença assertiva pronunciada como interrogativa, ou vice-versa**
- **Lentificação anormal do fluxo da fala**
- **Incongruências suprasegmentares não classificadas em outra parte:** anotação descritiva
- **Retratação:** releitura de trecho para correção de alguma(s) das incongruências acima (vide comentário sobre retratação vs. revisão, acima)

As incongruências listadas correspondem, neste estudo, ao que seriam as variáveis dependentes em um estudo experimental. Descreveremos, adiante, os aspectos do ambiente linguístico (onde as incongruências ocorreram) que serão analisados como se correspondessem às variáveis independentes. Nosso objetivo, portanto, é analisar as possíveis correlações entre incongruência e ambiente linguístico, partindo da hipótese de que o ambiente linguístico pode favorecer ou desfavorecer a ocorrência de incongruências na leitura.

Em suma, o conjunto de períodos lidos por cada falante foi analisado de acordo com a **Figura 2**:

Figura 2 – Fluxograma de análise do corpus



3.7 A validação das incongruências identificadas pelo pesquisador

A maioria das incongruências que escolhemos identificar e analisar são autoevidentes e não deixam muita margem para variações entre julgadores; mas há uma categoria de incongruência suprasegmentar que merece uma análise mais detida: as fronteiras prosódicas.

Segundo Cavalcante, Silva e Raso (2015, p. 141), “a segmentação em quebras prosódicas é pautada na percepção de falantes/ouvintes ‘competentes’ — no sentido de serem capazes de (de)codificar e interpretar pragmaticamente determinada língua”. Além disso, Raso, Mittmann e Mendes (2015) afirmam que não há dúvida de que os falantes conseguem perceber as fronteiras prosódicas e atribuir-lhes valor terminal ou não terminal com alto grau de concordância entre julgadores.

No entanto, em nosso estudo, não estaremos avaliando se julgadores conseguem identificar fronteiras prosódicas congruentes, mas se eles concordam na identificação de fronteiras prosódicas incongruentes, inseridas onde não deveriam estar, ou excluídas de onde deveriam estar (segundo um critério perceptivo, subjetivo, dependente da intuição de falantes nativos do PB).

Por essa razão, para saber se o julgamento do pesquisador com relação à existência de fronteiras prosódicas incongruentes encontra algum amparo na realidade linguística de outros falantes, selecionamos amostras aleatórias de incongruências de fronteira prosódica identificadas pelo pesquisador e as submetemos ao julgamento de quatro falantes do PB. Os dados de concordância e discordância obtidos foram submetidos ao teste Kappa de Cohen (ALBOUKADEL KASSAMBARA, 2021) e estão descritos em detalhe na seção Resultados.

3.8 Descrição do ambiente linguístico onde ocorreram (e não ocorreram) incongruências

Cada período incongruente (período em que ocorreu alguma incongruência) ou congruente (período em que não ocorreu nenhuma incongruência) foi analisado para caracterizar seu “ambiente linguístico”. Essa análise incluiu cada uma das seis características linguísticas descritas a seguir e teve por finalidade identificar uma eventual associação dessas características linguísticas com os “casos” (incongruência) e os “controles” (congruência).

3.8.1 Segmentação do trecho do discurso em que se insere a sentença que contém a incongruência, inspirado no esquema de segmentação de estruturas linguísticas em segmentos de discurso (*discourse segments* — DS) proposto por Grosz e Sidner (1986, p. 183).

Conforme mencionamos na Introdução, buscamos inspiração no modelo de Grosz e Sidner (1986) para segmentar nosso corpus e, adicionalmente, analisar se a relação entre segmentos do discurso poderia influenciar a ocorrência de incongruências. Como lidamos com fala lida, tomamos por referência a própria segmentação do texto escrito: os períodos. As unidades segmentares do discurso analisadas, portanto, são os períodos. Cada período foi classificado como incongruente ou congruente, conforme tenha apresentado, ou não, alguma incongruência.

A relação entre os períodos foi analisada do ponto de vista da continuidade temática. Portanto, cada período foi classificado — segundo sua relação temática com o período imediatamente anterior — em uma de duas possíveis categorias: 1) tematicamente contínuo; 2) tematicamente descontínuo. Em contraste com a proposta de Grosz e Sidner (1986), investigamos apenas as conexões temáticas entre períodos adjacentes, ou seja, não buscamos conexões temáticas entre períodos não adjacentes. Consideramos haver continuidade temática entre os períodos quando:

- o tópico do período atual (ou da oração principal do período atual) foi mencionado no período anterior, seja na posição de tópico, ou de foco
- o tópico do período atual é a primeira pessoa, pois a primeira pessoa (o falante) é sempre conhecida e é sempre tópico ou foco potencial em qualquer período do discurso
- o período atual foi introduzido por expressões que indicam conexão com o período anterior: “dessa forma”, “por isso”, “no entanto” etc

3.8.2 Determinação do grau de proeminência do ator do período, segundo o modelo AIS (*actor identification strategy*) proposto por Bornkessel-Schlesewsky e Schlewski (2014).

Como expusemos na Revisão da Literatura, consideramos o modelo “estratégia de identificação do ator” como sendo bastante plausível. Resolvemos, portanto, incluí-lo em nossa análise. Para cada período, identificamos o(s) atores(s) e o(s) classificamos de forma semelhante à proposta por Alday, Schlewski e Bornkessel-Schlesewsky (2013), de acordo com seu grau de proeminência. Nossa abordagem do grau de proeminência incluiu as seguintes categorias, dentro das quais cada ator foi

classificado (obs: evitamos o uso do termo “sujeito” em lugar de “ator” para evitar conflitos terminológicos e teóricos):

- Classificações genéricas:
 - voz passiva vs. voz ativa
 - sujeito desinencial (p.ex.: [eu] comi um hambúrguer)
 - sujeito coordenado (p.ex.: eu e meus amigos comemos hambúrgueres)
 - ✓ observação: orações subordinadas substantivas subjetivas não foram incluídas na classificação (p.ex.: controlar os nervos deixará você mais produtivo)
- Posição
 - antes vs. depois do verbo
- Pessoa
 - primeira pessoa vs. demais pessoas
- Humano
 - sim vs. não (p.ex.: o ladrão roubou a galinha vs. a galinha bicou o ladrão)
- Animado
 - sim vs. não (p. ex.: a galinha bicou o ladrão vs. penas voaram por todos os lados)
- Definido
 - sim vs. não (p. ex.: esta casa foi bem construída vs. casas bem construídas duram muito; ele caiu da escada vs. alguém caiu da escada)

3.8.3 Medição da extensão (em palavras) das sentenças onde ocorreram as incongruências.

Em linha com a escolha dos períodos (graficamente definidos, delimitados por ponto (“.”) no texto escrito) como unidades de segmentação do discurso, escolhemos as palavras gráficas (graficamente definidas, delimitadas por espaços em branco no texto escrito) como unidades de segmentação dos períodos (sentenças). Essa unidade de segmentação já reflete, apenas pelo fato de estar graficamente separada de outras unidades, uma realidade palpável na língua e na intuição dos falantes, independentemente de outras discussões conceituais que poderiam ser feitas e que, neste momento metodológico do estudo, desconsideramos (como, p.ex.: palavra gráfica vs. palavra fonológica vs. lexema vs. item lexical).

Portanto, todos os períodos do corpus (composto pelas 5h de fala lida) tiveram sua extensão aferida em “número de palavras”. Os resultados estão descritos e analisados na seção correspondente.

3.8.4 Determinação da taxa de elocução (em palavras por minuto) no trecho em que ocorreu a incongruência.

Tendo usado a “palavra gráfica” como unidade de segmentação dos períodos, nada mais natural do que aferir a taxa de elocução em palavras/minuto. Poderíamos ter optado, também, por aferir a taxa de elocução em sílabas fonológicas ou unidades V-V (unidades fonêmicas medidas entre vogais, de vogal a vogal) por unidade de tempo (via de regra, em segundos). Mas, como não estaremos comparando línguas diferentes, com números médios de sílabas/palavra diferentes (BARBOSA, 2019); e como estamos analisando um corpus com um grande número de palavras e, portanto, o número médio de sílabas por palavra por discurso por falante deve ser semelhante (apesar de não termos feito tal avaliação quantitativa); e como nosso objetivo é observar se a taxa de elocução é diferente entre o conjunto de períodos incongruentes em relação ao conjunto de períodos congruentes; por todos esses motivos, a segmentação do corpus em palavras gráficas não deve comprometer nossos resultados.

Outra possibilidade para aferir a velocidade com que os falantes emitem suas falas seria empregar, em vez da “taxa de elocução”, a “taxa de articulação”, que exclui da primeira os trechos silenciosos e, portanto, mede apenas o tempo de efetiva fonação. O uso da taxa de articulação poderia ser útil, pois excluiria os tempos de pausa incongruente, mas, por questões operacionais, não a utilizamos neste trabalho. Nas próximas seções (Resultados e Discussão), falaremos um pouco mais sobre essa escolha e suas consequências.

3.8.5 Análises das trocas de palavras

Nos casos em que houve troca de palavras, analisamos três aspectos: frequência da palavra na língua, classificação em categorias gramaticais, e semelhante entre as palavras trocadas. Essas análises são descritas abaixo:

- Foram pesquisadas as frequências relativas da palavra original (escrita) e da palavra inserida (lida) com base em um corpus do português brasileiro (SARDINHA; MOREIRA FILHO; ALAMBERT, 2019)
- As palavras (originais e inseridas) foram classificadas, segundo Payne (2006), em:
 - Gramaticais

- Lexicais
 - substantivo
 - adjetivo
 - advérbio
 - verbo
 - numeral
 - particípio
- As palavras (originais e inseridas) também foram classificadas, de acordo com sua semelhança fonológica/gráfica (paronímia), em:
 - Não semelhantes
 - Semelhantes
 - Com compartilhamento de lexema
 - Sem compartilhamento de lexema

Como já discutimos na seção Revisão de Literatura, os mecanismos de processamento de sentenças envolvem, entre outros, dois fenômenos cognitivos: *entrenchment* e predição.

O primeiro considera que eventos cognitivos mais frequentes estão mais bem “gravados na mente” e são mais disponíveis ao falante. Como já mencionamos, segundo Langacker (2017, p. 41), "*at some point, a unit is well enough established (deeply enough entrenched) that it constitutes ‘an event waiting to happen’: once initiated, it will run to completion if not deflected from this course*".

O segundo fenômeno, proposto, entre outros, por Pickering e Garrod (2013), acredita que nos antecipamos ao influxo linguístico e que prevemos (ou predizemos, ou pré-ouvimos, ou pré-lemos), entre outras coisas, palavras que ainda estão por surgir na cadeia linguística subsequente.

Palavras mais frequentes no corpus de uma língua específica devem ter, portanto, maior probabilidade de serem acessadas, interpretadas e pronunciadas; e, dessa forma, se houver, durante a leitura, troca de palavra, a palavra indevidamente inserida deve ser uma palavra mais frequente no corpus da língua.

Com relação à classificação das palavras originais e inseridas em gramaticais e lexicais (com suas subcategorias), nosso objetivo é observar se as trocas seguem algum padrão que se associe às categorias das palavras. Gostaríamos de saber se palavras lexicais têm uma frequência de troca

diferente de palavras gramaticais, ou se verbos são mais ou menos trocados do que substantivos, por exemplo.

Quanto ao último aspecto da análise das trocas de palavras, resolvemos investigar se a semelhança entre palavras influencia a escolha da palavra inserida. Para a análise de semelhança, partimos dos conceitos propostos por Basílio (2004, p. 12), que divide as palavras em unidades lexicais e unidades formais. As primeiras, também chamadas lexemas, carregariam a semântica básica da palavra; e as segundas seriam variações de forma da unidade lexical, como no caso das várias flexões de um determinado verbo, compartilhando um mesmo lexema. Ao contrário de Basílio (2004), no entanto, usaremos a mesma terminologia — lexema — para as palavras gramaticais. Consideraremos compartilhamento de lexema, por exemplo, o caso das palavras “de”, “da”, “disso” e “deste”, que compartilhariam o lexema/preposição “de”. Justificamos nossa postura com base na falta de consenso da literatura com relação a essa terminologia, como bem aponta Duarte (2004) ao listar o uso contraditório de termos como gramema, monema, semantema, morfema, entre outros.

3.8.6 Classificação hierárquica das palavras com relação à “predicação primária” (em períodos com incongruência por inserção/deleção inadequada de fronteiras prosódicas)

Por se tratar de uma abordagem metodológica pouco usual (ou, no conjunto da abordagem, talvez até inédita), este item merece um detalhamento especial. As fronteiras prosódicas que estamos analisando devem ser consideradas como fenômenos prosódicos de caráter linguístico cuja função principal é segmentar o fluxo da fala (BARBOSA, 2019). A partir dessas balizas prosódicas (fronteiras terminais e não terminais), podemos identificar determinadas unidades linguísticas (IZREEL *et al.*, 2020), delimitadas pelas fronteiras. Ou seja, as fronteiras e as unidades por elas delimitadas (ou as unidades e as fronteiras por elas induzidas) seriam duas faces de um mesmo fenômeno: a segmentação do fluxo da fala.

Segundo Izre’el (2018), as fronteiras não terminais e terminais delimitam o fluxo aparentemente contínuo da fala em unidades discursivas e sintáticas, conforme descrito na **Tabela 1**.

Tabela 1 - Unidades prosódicas vs. discursivas vs. sintáticas

Fronteira	Unidade discursiva	Unidade sintática
não terminal	módulo de informação	sintagma ou oração
terminal	enunciado	sentença/período

Ainda segundo Izre'el (2018, p. 1690), o **enunciado** seria “a unidade discursiva básica da fala espontânea”. Essa interpretação é a mesma defendida, entre outros, por Silva (2000, p. 271), quando nos lembra que a “*Language into Act Theory* é uma teoria pragmática corpus-driven da fala espontânea (...) [que] individualiza o **enunciado** como a unidade de referência da fala, sendo este definido como **a menor unidade linguística que possui autonomia pragmática e interpretabilidade em isolamento**. Essa teoria considera que a prosódia é responsável por delimitar não só o enunciado, mas também as ilocuções e a estrutura informacional (...)”.

Com esses conceitos em mente, resolvemos segmentar os textos escritos do nosso corpus (os períodos congruentes e os períodos em que houve incongruência por inserção/deleção de fronteiras prosódicas) segundo nossa própria interpretação do que poderia ser eventualmente considerado como:

- a) Enunciado
- b) Módulo de informação (segundo a **Tabela 1**)

O objetivo dessa segmentação foi tentar encontrar justificativas para as inserções/deleções de fronteiras feitas pelos leitores/falantes. Se uma determinada sequência de palavras dentro do período (um segmento de período) tivesse “autonomia pragmática” e “interpretabilidade em isolamento”, ela poderia ser (erroneamente) interpretada pelo leitor como correspondendo ao enunciado completo (o período completo). Neste caso, o leitor poderia inserir uma fronteira terminal que, no contexto real do período escrito, seria incongruente, pois o enunciado completo, o período real (escrito), incluía mais material linguístico do que o leitor imaginou.

Outra possibilidade (item “b”, acima) são os módulos de informação: o leitor poderia acreditar que determinado módulo chegou ao final e, assim, inserir uma fronteira não terminal. Em ambos os casos, o raciocínio é que o leitor, no correr do texto escrito, estando preso ao momento da leitura (palavra por palavra, sem saber exatamente quais e quantas seriam as próximas palavras do período, ou qual a mensagem completa pretendida pelo período), poderia, eventualmente, imaginar que determinada transição entre palavras corresponderia ao limite de uma unidade discursiva e, assim, inserir uma fronteira prosódica que fosse incongruente com o período real, escrito.

Do ponto de vista metodológico, o procedimento adotado pelo pesquisador foi o seguinte:

- Leitura dos textos escritos originais
- Segmentação dos períodos conforme sua percepção de:
 - autonomia pragmática e interpretabilidade em isolamento (vide acima) do segmento lido
 - possibilidade de completar um “módulo de informação” sem necessidade da(s) palavra(s) subsequente(s)

Ressalte-se que a análise foi feita levando em conta uma característica fundamental da língua (escrita, falada...), que é a serialidade. Portanto, cada palavra foi classificada sempre com relação à palavra anterior, e nunca considerando a(s) palavra(s) seguinte(s), buscando simular a experiência do leitor/orador. Exemplos concretos, retirados do corpus, são fornecidos abaixo.

Analisando o período “Essa corporação — cujo aniversário de 159 anos ocorreu no último dia 28 — carrega a um só tempo as marcas da tradição e do avanço.”, segmentamo-lo da seguinte forma (note-se que condensamos a imagem do período para caber em uma única linha e facilitar a compreensão da segmentação):

1) Essa corporação		carrega	as marcas da tradição	
2)	- cujo aniversário	ocorreu no último dia 28 -	a um só tempo	e do avanço.
3)	de 159 anos			

- Linha “1”: corresponde à hierarquia principal do período, um segmento que poderia ser interpretado como um enunciado; neste trabalho, resolvemos chamá-lo de “predicação primária”
- Demais linhas: palavras/frases/orações cuja omissão não prejudica a integridade da predicação primária, nem da linha imediatamente superior:
 - Linha “2”: palavras/frases/orações cuja omissão não prejudica a integridade (sintática/discursiva/pragmática) da linha “1”
 - Linha “3”: palavras/frases/orações cuja omissão não prejudica a integridade (sintática/discursiva/pragmática) da linha “2”

De fato, essa hierarquia pode ser percebida relendo o período de acordo com a segmentação proposta:

- Predicação primária (linha 1): “Essa corporação carrega as marcas da tradição”
- Acréscimo da segunda hierarquia (**linha 2**, em negrito) à predicação primária: “Essa corporação — **cujo aniversário ocorreu no último dia 28** — carrega **a um só tempo** as marcas da tradição **e do avanço.**”
- Acréscimo da terceira hierarquia (**linha 3**, em negrito) às anteriores: “Essa corporação — cujo aniversário **de 159 anos** ocorreu no último dia 28 — carrega a um só tempo as marcas da tradição e do avanço.”

O exemplo a seguir ilustra alguns aspectos adicionais dessa segmentação:

- | | | | |
|----|-------------|--------------|----------------|
| 1) | eu não vejo | | solução |
| 2) | Só que | nas emendas | uma para isso. |
| 3) | | ao orçamento | |

A predicação primária seria: “eu não vejo solução”. É a menor unidade linguística interpretável em isolamento e poderia ser considerada como um enunciado. Seus elementos são obrigatórios. O acréscimo dos elementos das linhas 2 e 3 é “opcional”, do ponto de vista da interpretabilidade (apesar de não ser opcional do ponto de vista da mensagem pretendida por quem criou o período).

Observe-se que os elementos que não pertencem à predicação primária (qualquer elemento em uma linha diferente de “1”) não têm conexão recíproca uns com os outros, exceto quando são adjacentes. Por exemplo, é possível adicionar à predicação primária qualquer elemento da linha 2, sem adicionar os demais:

- “**só que** eu não vejo solução”
- “eu não vejo solução **para isso**”
- “eu não vejo **nas emendas uma** solução”
- “**só que** eu não vejo **uma** solução”
- “**só que** eu não vejo **nas emendas uma** solução **para isso**”
- e assim por diante...

Com relação à linha 3, seus elementos também só estão ligados aos elementos hierarquicamente superiores (linha 2) quando lhes são adjacentes e subsequentes. Isso significa que a inserção dos elementos da linha 3 depende da inserção dos elementos da linha 2 que lhes sejam adjacentes e antecedentes, como ilustrado abaixo:

- “não vejo [nas emendas] linha 2 [**ao orçamento**] linha 3 solução”

Naturalmente, não se trata de uma ciência exata, mas apenas de uma tentativa de identificar as possibilidades de agrupar elementos linguísticos dos períodos escritos em unidades menores, respeitando a predicação primária. Classificações alternativas seriam possíveis, como, por exemplo, excluir o pronome “eu” da predicação primária: “não vejo solução”. Mas, como nos propusemos a classificar uma grande quantidade de material linguístico, acreditamos que, no geral, a classificação aproxima-se, estatisticamente, do que seria o correspondente à nossa proposta inicial.

Para finalizar esta exposição e ressaltar o critério de serialidade nessa classificação hierárquica, vejamos mais um período:

1) Matsubara também foi o responsável por atrair japoneses para cidades

2)

onde a colônia envolveu-se no cultivo de borracha.

3)

no atual Estado do Mato Grosso,

e produção

Note-se que, em construções coordenadas, via de regra é possível ignorar os elementos posteriores, como a oração “...onde a colônia envolveu-se no cultivo...” e o constituinte “produção”, coordenado em “cultivo e produção”. A oração “a colônia envolveu-se no cultivo de borracha” poderia, caso surgisse antes da outra oração (“Matsubara (...) foi o responsável...”), ser considerada como a predicação primária, mas, neste caso específico, como veio depois, não.

Pode-se argumentar que esse sistema de classificação se baseia no conceito de argumentos vs. adjuntos. Por ora, não vamos aprofundar a discussão, que deixaremos para uma seção mais apropriada, pós-resultados (Discussão). Apenas contra-argumentamos que existe grande discussão na literatura em torno desses termos e não há consenso quanto à abrangência ou mesmo quanto à existência real do que tais termos querem descrever. Então, neste momento, apenas explicamos os critérios de classificação “hierárquica” que utilizamos e os resumimos da seguinte forma: obrigatoriedade vs. opcionalidade de cada palavra (com relação à palavra antecedente), respeitada a serialidade da língua.

Nossas hipóteses com relação a esta análise são as seguintes:

- As predicções primárias estarão preservadas (não serão fragmentadas)
- Não haverá inserção de fronteira terminal antes de completar a predicação primária
- Haverá mais incongruências de fronteira depois da predicação primária

- As inserções incongruentes de fronteira ocorrerão, majoritariamente, nos desnivelamentos (quebras de linha) da hierarquia argumental

3.9 Aspectos complementares e coda metodológica

As análises estatísticas foram realizadas no programa RStudio (2020). As referências bibliográficas foram formatadas no sistema MORE (MORE, 2013). E, não obstante as diversas informações consideradas, discutidas e ponderadas até aqui, devemos lembrar um conselho que Darcy Ribeiro (RIBEIRO, 1995), a propósito de um de seus trabalhos, confessa ter recebido de Claude Lévi-Strauss: "teorias vivem pouco, dados e descrições nunca morrem". Aos dados, então.

4. RESULTADOS

Os resultados deste trabalho estão descritos e analisados da seguinte forma:

- 4.1 Falantes, corpus, incongruências *lato sensu* e padrões respiratórios
- 4.2 Incongruências segmentares
- 4.3 Incongruências suprasegmentares
- 4.4 Predicação primária: palavras/constituintes obrigatórios vs. opcionais

4.1 Falantes, corpus e incongruências *lato sensu*

Esta seção organiza-se da seguinte forma:

- 4.1.1 Os falantes
- 4.1.2 O corpus
- 4.1.3 As incongruências *lato sensu*
 - 4.1.3.1 Os períodos congruentes e incongruentes
 - 4.1.3.2 Congruência vs. extensão dos períodos
 - 4.1.3.3 Congruência vs. taxas de elocução
 - 4.1.3.4 Congruência vs. idade
 - 4.1.3.5 Congruência vs. sexo
- 4.1.4 Os padrões respiratórios
- 4.1.5 Resumo da seção
- 4.1.6 Prólogo das próximas seções

4.1.1 Os falantes

Analisamos 1610 textos escritos produzidos entre os anos de 2017 e 2018 e identificamos 39 falantes que leram publicamente pelo menos um desses textos. Alguns falantes leram mais de um texto e, nesses casos, escolhemos apenas o texto com a maior quantidade de material lido. Dessa forma, selecionamos apenas um texto por falante, obtendo uma amostra de 39 textos lidos que, acreditamos, reflete de forma mais equilibrada o universo de falantes. A **Tabela 2** descreve as características demográficas desses falantes.

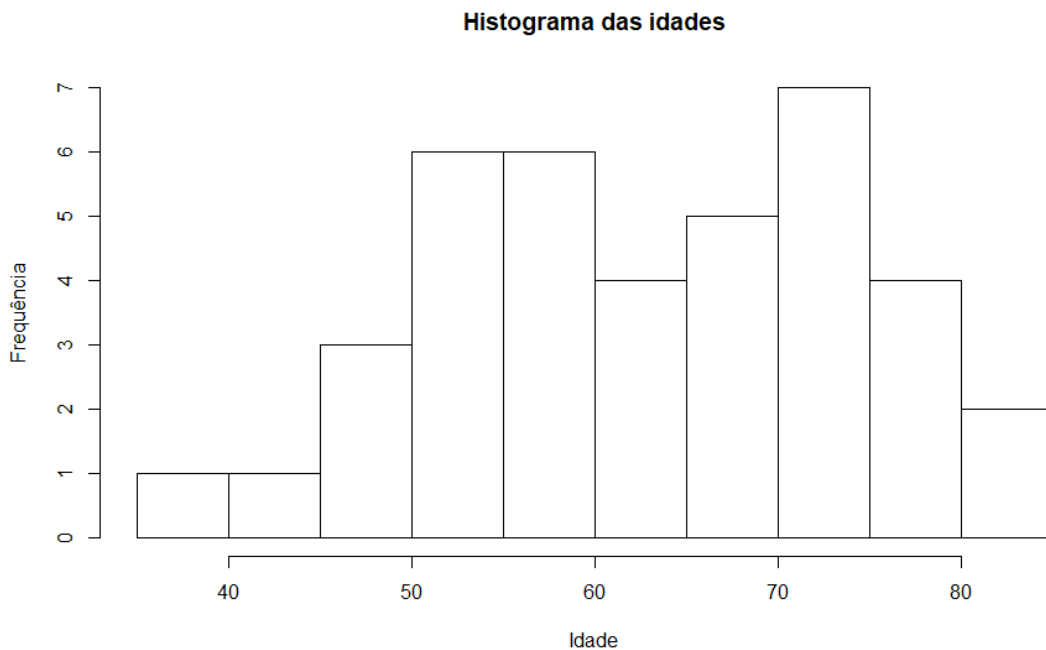
Tabela 2 - Características demográficas dos falantes no momento da fala

Falante	Idade	Sexo	Naturalidade	Escolaridade
1	74	M	SP	SUPERIOR
2	69	M	PE	SUPERIOR
3	75	M	AL	SUPERIOR
4	54	M	PB	SUPERIOR
5	50	M	PI	SUPERIOR
6	61	M	SC	SUPERIOR
7	54	M	SE	SUPERIOR
8	56	M	PA	SUPERIOR
9	54	M	SP	SUPERIOR
10	75	M	CE	SUPERIOR
11	65	M	CE	SUPERIOR
12	71	M	RJ	SUPERIOR
13	70	M	RN	SUPERIOR
14	50	M	AL	SUPERIOR
15	39	M	AC	SUPERIOR
16	58	M	GO	SUPERIOR
17	64	F	GO	SUPERIOR
18	81	M	MA	SUPERIOR
19	85	M	PB	SUPERIOR
20	48	M	RN	SUPERIOR
21	75	M	RS	SUPERIOR
22	61	F	BA	SUPERIOR
23	72	F	GO	SUPERIOR
24	43	F	MS	MÉDIO
25	73	F	SP	SUPERIOR
26	59	M	SP	SUPERIOR
27	68	M	RS	FUNDAMENTAL
28	67	M	PA	MÉDIO
29	78	M	MS	SUPERIOR
30	75	M	PB	SUPERIOR
31	52	M	RJ	SUPERIOR
32	70	F	MG	SUPERIOR
33	55	M	RR	SUPERIOR
34	76	M	MG	SUPERIOR
35	60	M	RR	SUPERIOR
36	62	M	SC	SUPERIOR
37	56	F	SC	SUPERIOR
38	60	M	MT	SUPERIOR
39	49	M	GO	SUPERIOR
Média	63
DP	11

Algumas características da amostra de falantes podem ser destacadas:

- Há falantes nascidos em 20 estados brasileiros, estando excluídos apenas seis estados — Amapá, Amazonas, Espírito Santo, Paraná, Rondônia, Tocantins — e o Distrito Federal.
- As mulheres compõem 18% (sete indivíduos) da amostra e os homens, 82% (32 indivíduos).
- Apenas três falantes não possuem Ensino Superior, sendo dois homens e uma mulher.
- A faixa etária estende-se de 39 a 84 anos, com uma média de 63 anos (vide **Figura 3**) e distribuição próxima à normal.

Figura 3 – Histograma de distribuição das idades



4.1.2 O corpus

A duração total dos arquivos de áudio/vídeo analisados foi de 8 horas e 28 minutos. Desse total, foram identificados e extraídos 5 horas de áudio/vídeo contendo falas lidas. As falas foram consideradas “lidas” apenas quando corresponderam aos textos escritos disponíveis para o pesquisador. Trechos de falas que foram lidas com base em textos não acessíveis para o pesquisador não foram considerados, computados ou analisados como “falas lidas”.

O tempo **fala total** foi computado com base na duração total dos pronunciamentos, independentemente de terem sido lidos. Os tempos de **fala lida**, por sua vez, foram computados a partir da soma dos tempos de duração dos períodos. Os tempos de fala lida, portanto, excluem interrupções ocorridas no intervalo entre períodos, mas incluem interrupções (pausas, hesitações etc) ocorridas dentro dos períodos. A unidade (segmento mínimo) de análise para os trechos de fala lida congruentes ou incongruentes, portanto, foi o período.

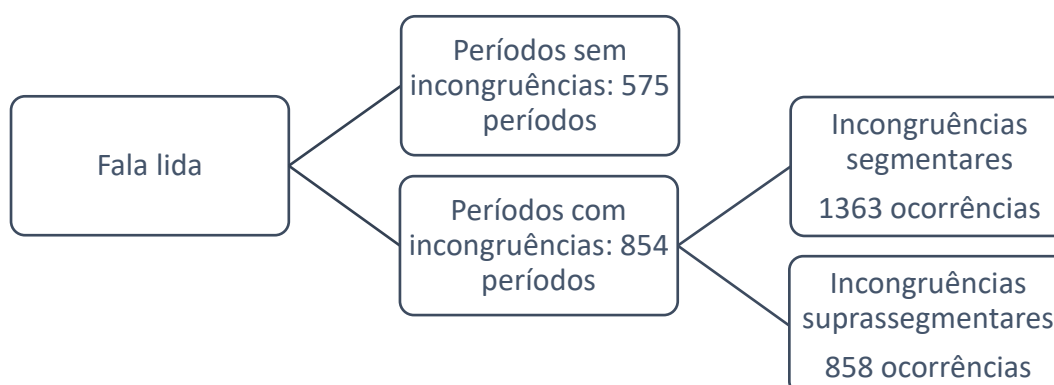
4.1.3 As incongruências

4.1.3.1 Períodos congruentes e incongruentes

A análise da fala lida obedeceu aos seguintes passos, descritos esquematicamente na **Figura 4**:

1. Os trechos dos arquivos de áudio/vídeo em que houve leitura foram segmentados em períodos.
2. Os períodos foram classificados como **congruentes** ou **incongruentes**.
3. Os períodos incongruentes tiveram suas incongruências identificadas e classificadas em **segmentares** ou **suprasegmentares**.
4. Cada incongruência foi anotada seguindo uma classificação específica dentro da respectiva categoria (segmentar ou suprasegmentar), conforme descrito na Metodologia.

Figura 4 – Fluxograma da análise do corpus



A **Tabela 3** descreve, para cada falante, os tempos de fala total, de fala lida, de fala congruente (soma dos períodos lidos com congruência) e de fala incongruente (soma dos períodos lidos com alguma incongruência).

Tabela 3 - Duração das falas (horas, minutos e segundos - hh:mm:ss)

Falante	Fala total	Fala lida	Fala congruente	Fala incongruente
1	00:17:00	00:15:34	00:04:07	00:11:26
2	00:10:00	00:02:26	00:01:26	00:01:00
3	00:19:48	00:08:16	00:01:21	00:06:54
4	00:07:00	00:00:18	00:00:00	00:00:18
5	00:08:10	00:03:30	00:00:09	00:03:20
6	00:15:00	00:09:09	00:01:49	00:07:20
7	00:12:35	00:09:11	00:02:21	00:06:49
8	00:18:57	00:10:44	00:02:58	00:07:45
9	00:08:39	00:06:06	00:03:10	00:02:56
10	00:16:54	00:12:02	00:01:52	00:10:10
11	00:19:12	00:07:39	00:03:45	00:03:54
12	00:25:26	00:13:51	00:08:11	00:05:40
13	00:12:01	00:06:52	00:04:15	00:02:37
14	00:12:18	00:09:33	00:03:06	00:06:28
15	00:11:50	00:06:55	00:00:41	00:06:14
16	00:12:36	00:07:10	00:02:10	00:05:00
17	00:08:06	00:06:18	00:02:12	00:04:06
18	00:09:33	00:08:22	00:00:37	00:07:46
19	00:07:00	00:05:40	00:00:36	00:05:04
20	00:13:33	00:06:56	00:02:47	00:04:09
21	00:16:04	00:12:23	00:05:03	00:07:20
22	00:06:52	00:00:57	00:00:17	00:00:40
23	00:15:55	00:08:48	00:03:58	00:04:50
24	00:09:18	00:02:13	00:01:15	00:00:58
25	00:11:39	00:08:37	00:05:46	00:02:50
26	00:10:32	00:06:36	00:00:58	00:05:38
27	00:24:55	00:19:31	00:08:57	00:10:34
28	00:15:35	00:08:02	00:00:14	00:07:48
29	00:12:45	00:11:18	00:03:12	00:08:06
30	00:16:09	00:08:41	00:02:54	00:05:46
31	00:08:18	00:03:10	00:00:52	00:02:18
32	00:18:34	00:08:53	00:03:13	00:05:41
33	00:06:30	00:05:39	00:00:28	00:05:11
34	00:10:20	00:07:14	00:03:41	00:03:33
35	00:12:00	00:08:14	00:01:17	00:06:57
36	00:10:02	00:08:15	00:01:23	00:06:52
37	00:13:20	00:03:07	00:01:04	00:02:03
38	00:13:48	00:07:53	00:02:19	00:05:34
39	00:10:45	00:03:24	00:00:24	00:02:59
Mediana	00:12:18	00:07:53	00:02:10	00:05:34
Total	08:28:59	04:59:24	01:34:50	03:24:34

Pode-se observar que (considerando as medianas), de um tempo de fala total de 12 minutos, o tempo de fala lida por falante foi de quase 8 minutos, sendo pouco mais de 2 minutos de fala congruente e 5 minutos e meio de fala incongruente. Como os tempos de fala congruente e incongruente correspondem, respectivamente, à soma da duração dos períodos congruentes e incongruentes, é importante discriminar o número de períodos congruentes e incongruentes identificados. Esses dados estão descritos na **Tabela 5**, segregados por falante.

Tabela 4 - Número de períodos com e sem incongruências

Falante	Sem incongruências	Com incongruências	Total
1	26	50	76
2	10	7	17
3	8	26	34
4	0	1	1
5	1	13	14
6	11	34	45
7	17	36	53
8	16	28	44
9	13	14	27
10	10	29	39
11	13	9	22
12	34	20	54
13	16	9	25
14	27	37	64
15	5	25	30
16	25	33	58
17	11	15	26
18	3	29	32
19	8	22	30
20	25	22	47
21	30	34	64
22	2	2	4
23	20	17	37
24	8	4	12
25	37	13	50
26	9	24	33
27	58	47	105
28	1	26	27
29	21	34	55
30	13	19	32
31	8	11	19
32	21	24	45
33	3	19	22
34	14	12	26
35	18	37	55
36	9	26	35
37	7	9	16
38	13	21	34
39	4	16	20
Mediana	13	22	35
Total	575	854	1429

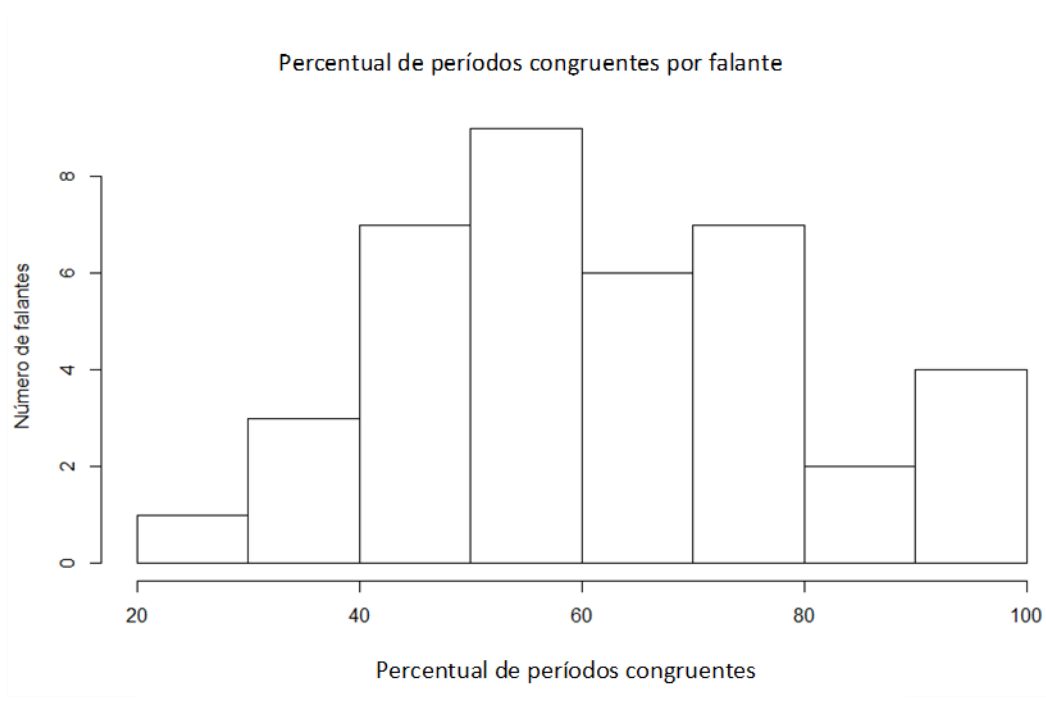
Observa-se que todos os falantes produziram incongruências durante a leitura. A maioria deles (29 falantes) produziu mais períodos incongruentes do que congruentes. A **Tabela 5** exibe o percentual de períodos incongruentes por falante, mostrando que 59% (mediana) dos períodos foram lidos com alguma incongruência.

Tabela 5 - Percentual de períodos com e sem incongruências por falante, ordenados pelo percentual de incongruências

Falante	Sem incongruência	Com incongruência
4	0%	100%
28	4%	96%
5	7%	93%
18	9%	91%
33	14%	86%
15	17%	83%
39	20%	80%
3	24%	76%
6	24%	76%
10	26%	74%
36	26%	74%
19	27%	73%
26	27%	73%
7	32%	68%
35	33%	67%
1	34%	66%
8	36%	64%
29	38%	62%
38	38%	62%
30	41%	59%
31	42%	58%
14	42%	58%
17	42%	58%
16	43%	57%
37	44%	56%
32	47%	53%
21	47%	53%
9	48%	52%
22	50%	50%
20	53%	47%
34	54%	46%
23	54%	46%
27	55%	45%
2	59%	41%
11	59%	41%
12	63%	37%
13	64%	36%
24	67%	33%
25	74%	26%
Mediana	41%	59%

Para observar a distribuição do percentual de períodos incongruentes por falante, construímos o histograma da **Figura 5**.

Figura 5 – Percentual de períodos incongruentes por falante



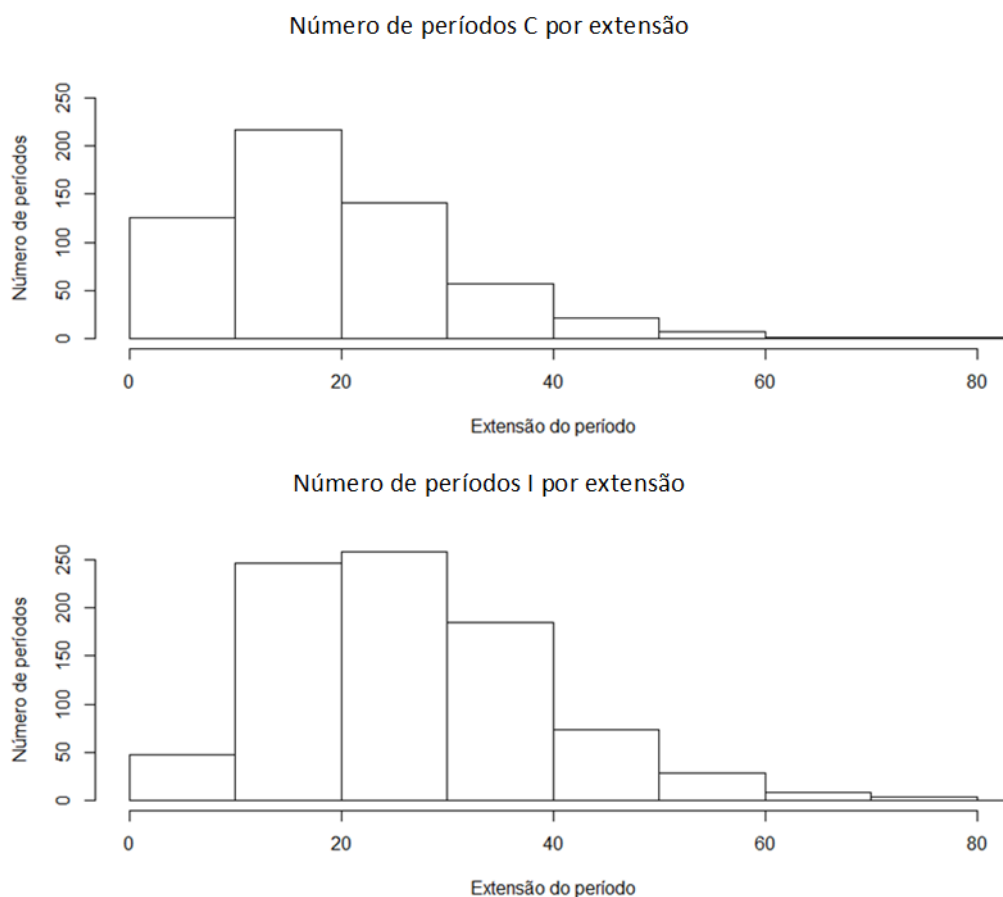
Este histograma (**Figura 5**) apresenta uma distribuição próxima da normal, exceto pelo acúmulo de falantes no intervalo de 90-100%. Parte desse acúmulo deve-se à inclusão do falante 4, um *outlier*, que leu apenas um período, incongruente, obtendo um percentual de incongruências (não representativo) de 100%.

4.1.3.2 Congruência vs. extensão dos períodos

Os períodos analisados são compostos por um número variável de palavras. Se considerarmos o universo total de períodos lidos (1429 períodos), podemos construir histogramas de frequência para

as diferentes extensões dos períodos, dividindo-os em incongruentes (I) e congruentes (C). A **Figura 6** mostra esses histogramas.

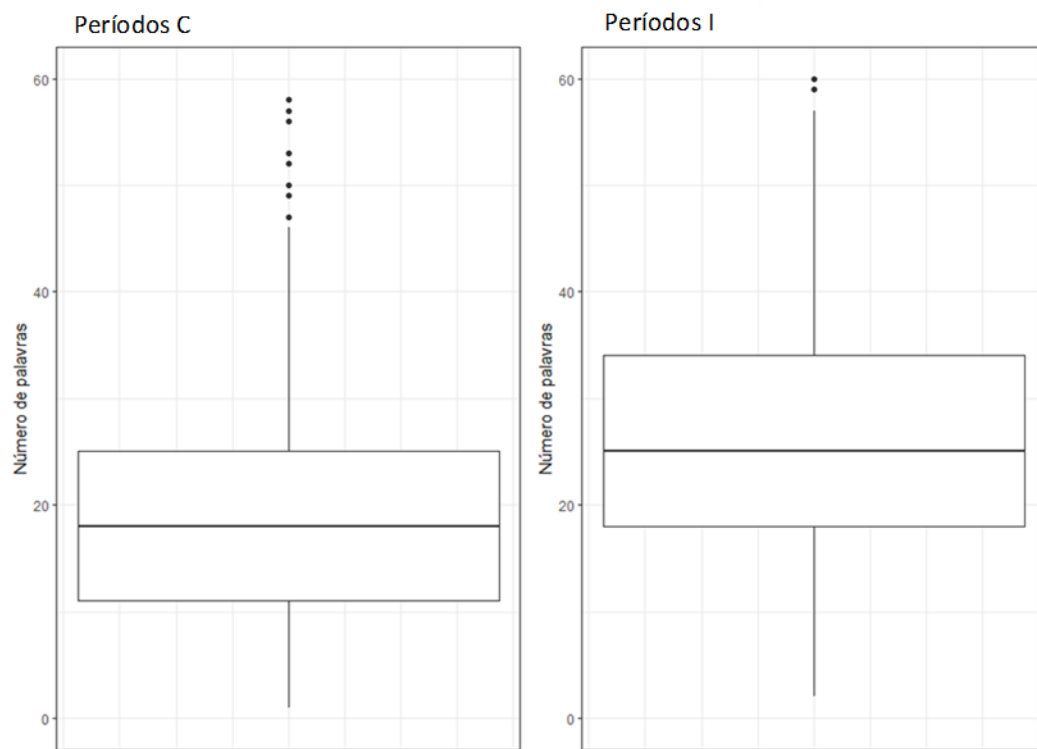
Figura 6 – Número de períodos por extensão (em palavras)



É possível observar que a distribuição dos períodos I (incongruentes) se concentra em uma extensão acima de 20 palavras, enquanto que a dos períodos C (congruentes) se concentra abaixo de 20 palavras. O teste de Wilcoxon (*rank sum test*) confirma essa diferença de extensão em favor dos períodos I, com um *p-value* menor do que $2,2e-16$. Ou seja, os períodos incongruentes são mais extensos (em número de palavras) do que os períodos congruentes.

Os diagramas de caixa da **Figura 7** ilustram melhor a diferença de extensão entre períodos C e I. Para facilitar a comparação, foram excluídos dos gráficos algumas extensões extremas (*outliers*), acima de 60 palavras.

Figura 7 – Extensão (em número de palavras) dos períodos C e I



Esses dados de extensão média de períodos I e C também podem ser discriminados por falante. A **Tabela 6** mostra o total de palavras em cada tipo de período e a **Tabela 7** calcula o número médio de palavras por período (extensão dos períodos), por falante (excluído o falante 4, que não produziu período congruente).

Tabela 6 - Número de palavras e de períodos por tipo de período

Falante	Total de palavras em períodos C	Períodos C	Total de palavras em períodos I	Períodos I	Total de palavras	Total de períodos
1	499	26	1281	50	1780	76
2	152	10	107	7	259	17
3	135	8	660	26	795	34
5	28	1	493	13	521	14
6	174	11	727	34	901	45
7	286	17	818	36	1104	53
8	345	16	870	28	1215	44
9	354	13	336	14	690	27
10	173	10	870	29	1043	39
11	340	13	338	9	678	22
12	996	34	611	20	1607	54
13	350	16	227	9	577	25
14	393	27	831	37	1224	64
15	90	5	698	25	788	30
16	287	25	641	33	928	58
17	311	11	602	15	913	26
18	68	3	738	29	806	32
19	70	8	520	22	590	30
20	420	25	606	22	1026	47
21	578	30	808	34	1386	64
22	33	2	82	2	115	4
23	434	20	534	17	968	37
24	157	8	129	4	286	12
25	848	37	401	13	1249	50
26	133	9	786	24	919	33
27	1056	58	1163	47	2219	105
28	21	1	745	26	766	27
29	343	21	886	34	1229	55
30	370	13	702	19	1072	32
31	147	8	320	11	467	19
32	327	21	590	24	917	45
33	62	3	576	19	638	22
34	414	14	391	12	805	26
35	157	18	726	37	883	55
36	155	9	790	26	945	35
37	136	7	219	9	355	16
38	280	13	623	21	903	34
39	46	4	368	16	414	20
Média	294	15	600	22	894	38
Mediana	283	13	617	22	902	34

Tabela 7 - Palavras por tipo de período

Falante	Palavras/período C	Palavras/período I
1	19	26
2	15	15
3	17	25
5	28	38
6	16	21
7	17	23
8	22	31
9	27	24
10	17	30
11	26	38
12	29	31
13	22	25
14	15	22
15	18	28
16	11	19
17	28	40
18	23	25
19	9	24
20	17	28
21	19	24
22	17	41
23	22	31
24	20	32
25	23	31
26	15	33
27	18	25
28	21	29
29	16	26
30	28	37
31	18	29
32	16	25
33	21	30
34	30	33
35	9	20
36	17	30
37	19	24
38	22	30
39	12	23
Média	19	28
Mediana	19	28

Os histogramas a seguir mostram a distribuição do número de falantes por média de extensão dos períodos C (**Figura 8**) e I (**Figura 9**). Perceba-se que estes gráficos são diferentes dos histogramas da **Figura 4**, que mostram a distribuição geral do número de períodos por extensão dos períodos C e I.

Figura 8 – Média (por falante) do número de palavras por período congruente (C)

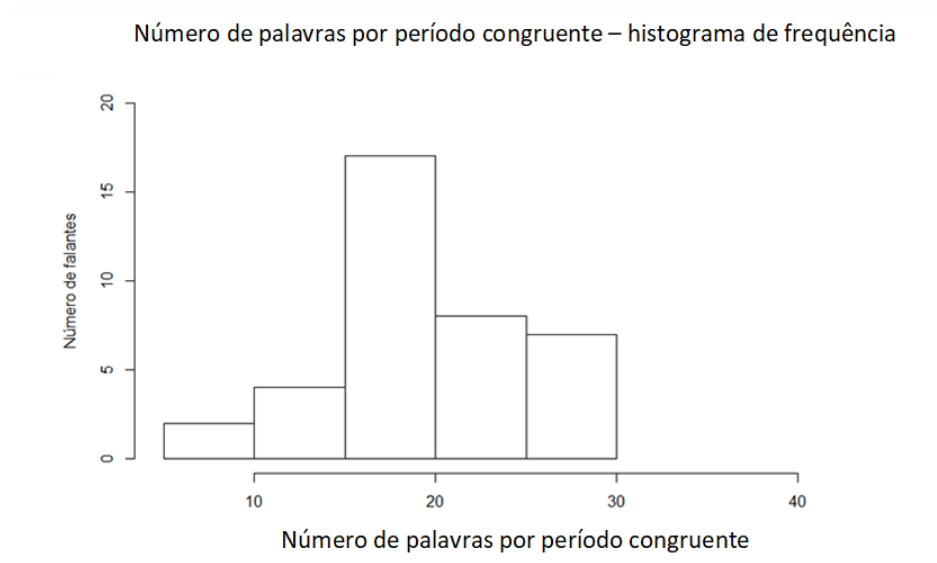
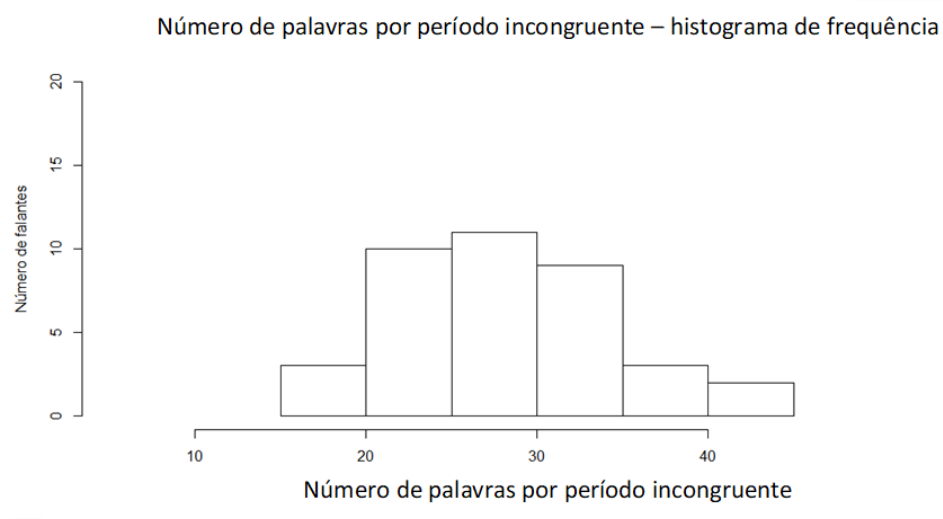


Figura 9 – Média (por falante) do número de palavras por período incongruente (I)



O teste t para amostras pareadas confirma que os períodos I têm de 7 a 10 palavras (intervalo de confiança) a mais do que os períodos C, com um *p-value* de 2,808e-13 (bastante significativo). Esse é um achado que não surpreende, pois parece natural que períodos mais extensos tenham mais incongruências, uma vez que cada palavra adicional acrescenta uma possibilidade a mais de incongruência. Resta analisar se a incidência de incongruências correlaciona-se linearmente com o aumento de palavras (cada palavra a mais acrescenta a mesma probabilidade de ocorrer incongruência), ou se essa correlação é de natureza não linear (cada palavra a mais acrescenta uma probabilidade ainda maior — ou menor — de ocorrer incongruência).

Para visualizar a correlação entre tamanho do período e incidência de incongruências, elaboramos três gráficos de correlação/dispersão. Optamos por incluir todos os períodos lidos (congruentes e incongruentes) e correlacionar o número de incongruências de cada tipo (segmentar ou suprasegmentar) com o tamanho do período. Nesses gráficos, o eixo “x” representa a extensão dos períodos em número de palavras e o eixo “y” representa, respectivamente:

- **Figura 10:** número médio de incongruências (de qualquer tipo) em períodos de tamanho “x”
- **Figura 11:** número médio de incongruências do tipo segmentar em períodos de tamanho “x”
- **Figura 12:** número médio de incongruências do tipo suprasegmentar em períodos de tamanho “x”

Figura 10 – Número médio de incongruências vs. tamanho do período: plotagem e linha de melhor ajuste

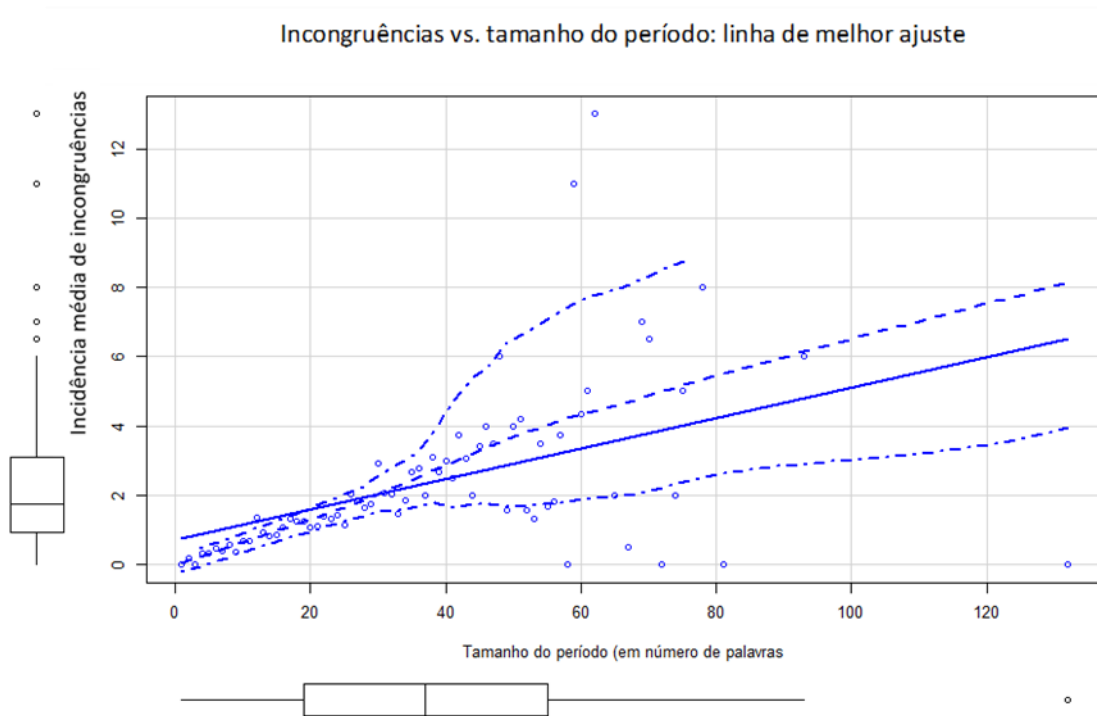


Figura 11 – Número médio de incongruências segmentares vs. tamanho do período: plotagem e linha de melhor ajuste

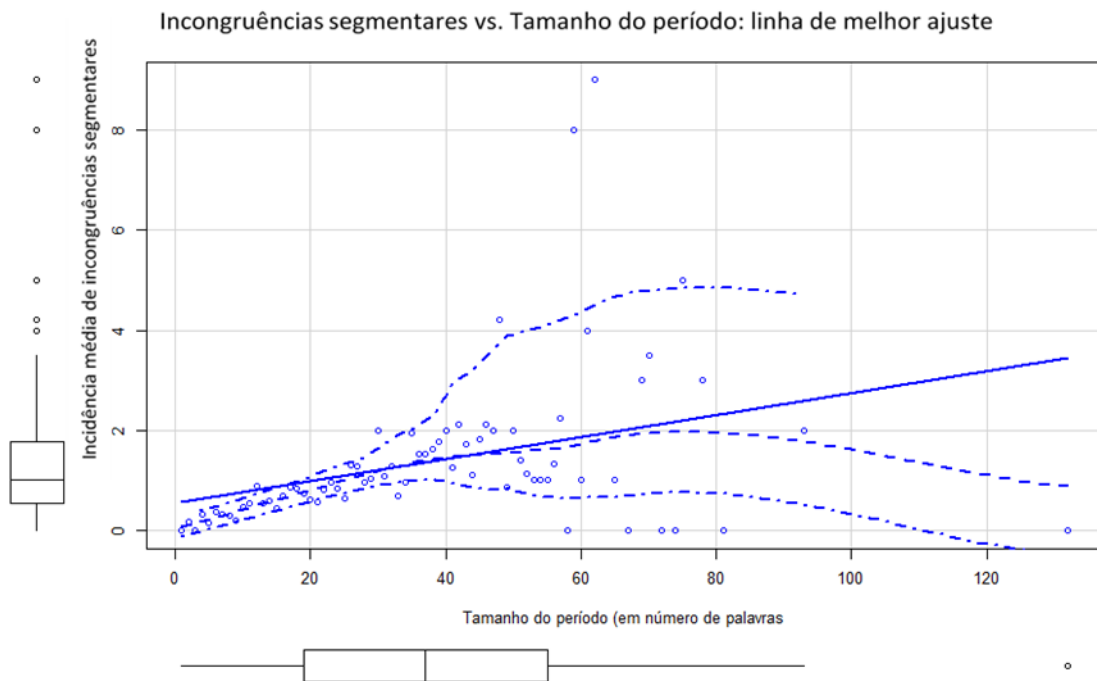
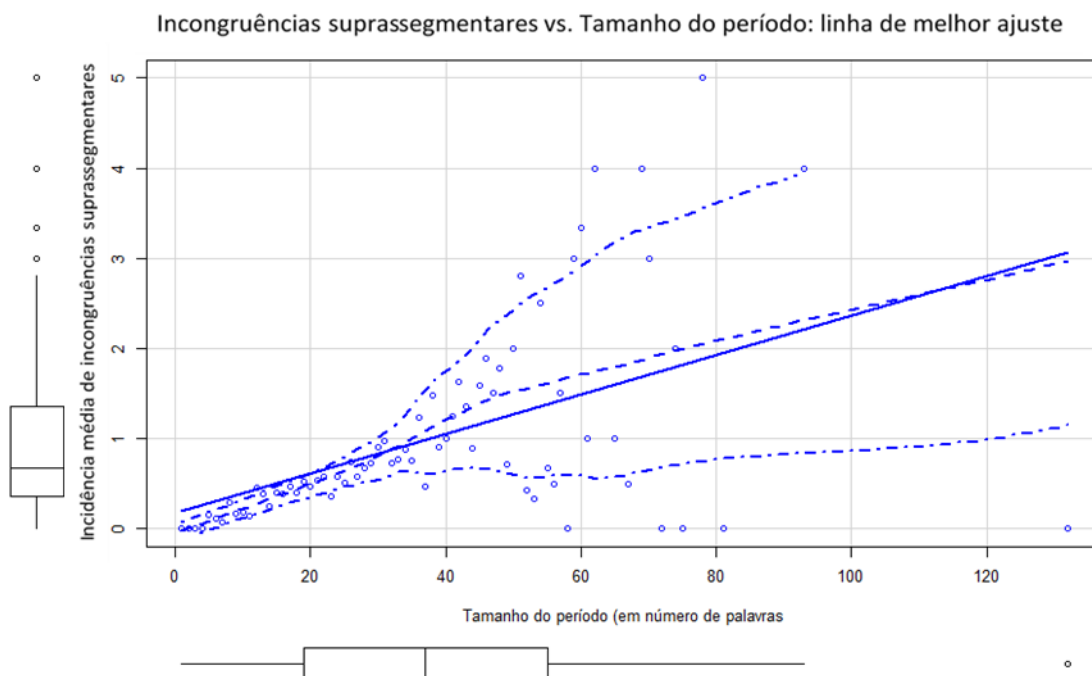
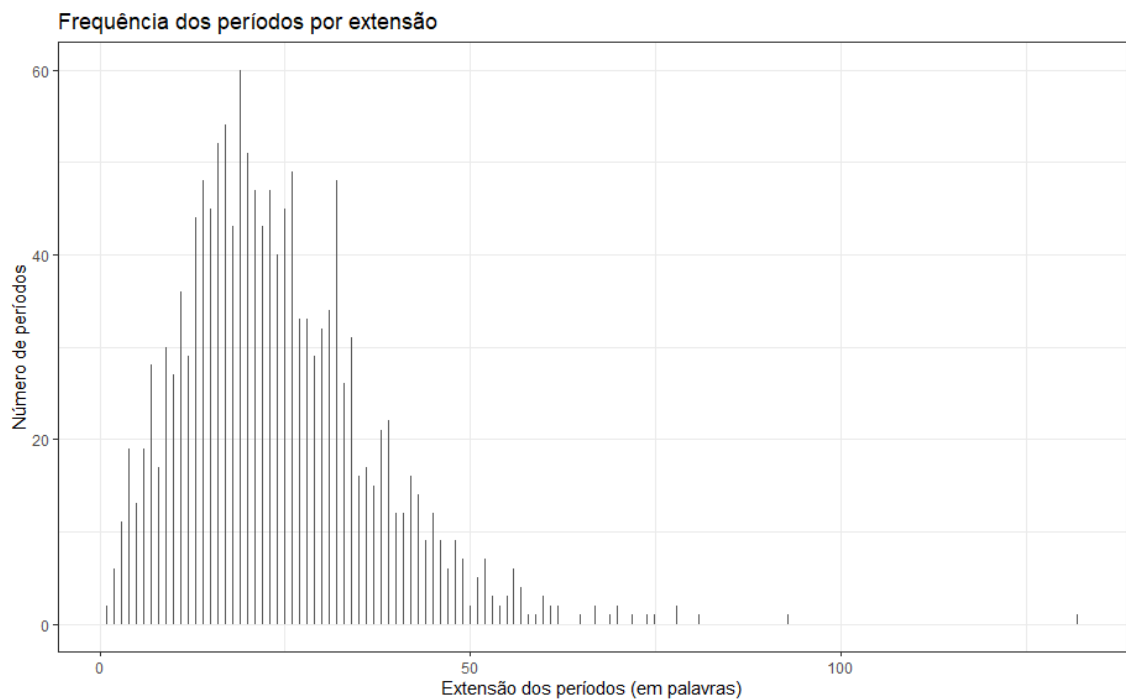


Figura 12– Número médio de incongruências suprasegmentares vs. tamanho do período: plotagem e linha de melhor ajuste



Os gráficos acima mostram grande dispersão nos valores médios de incongruência nos períodos com mais de 45-50 palavras, o que se deve (provavelmente) ao fato de que a amostra é pouco representativa para períodos com essa extensão (são períodos menos frequentes) e, portanto, as médias são geradas com base em poucos valores ou, até mesmo, em apenas um valor. Uma visão geral da distribuição por extensão de todos os períodos lidos — evidenciando a baixa incidência de períodos com mais de 50 palavras — pode ser visto na **Figura 13**.

Figura 13 – Número de períodos por extensão (em palavras)



Apesar da dispersão acima de 45-50 palavras, o comportamento do gráfico no intervalo entre 0 e 40 palavras sugere uma correlação linear entre tamanho dos períodos e incidência de incongruências. Ou seja, em média, cada palavra a mais adiciona um risco semelhante a mais de ocorrência de incongruência no período.

4.1.3.3 Congruência vs. taxa de elocução

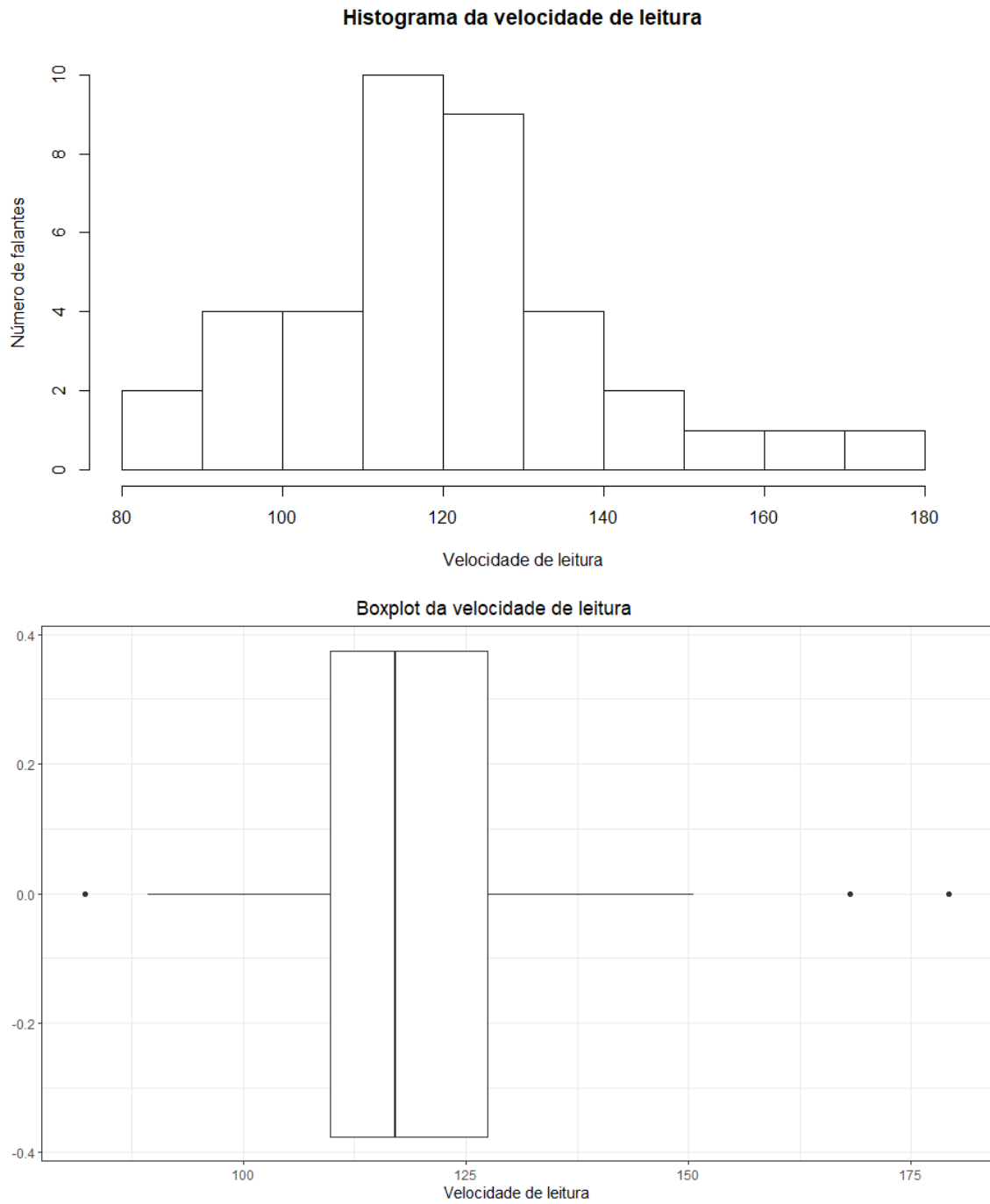
Outra variável analisada foi a taxa de elocução, quantificada em palavras faladas por minuto. A **Tabela 8** descreve a taxa de elocução (v média) por falante, discriminando-a por tipo de período (vD = velocidade incongruente; vF = velocidade congruente). O falante 4 foi excluído da análise, pois leu apenas um período I e, portanto, não produziu dados de taxa de elocução para período C.

Tabela 8 - Taxa de elocução (v) em palavras/minuto, em ordem crescente

Falante	v média	vF	vD
13	84	82	87
10	87	93	86
11	89	91	87
28	95	89	96
3	96	99	96
18	96	111	95
6	99	96	99
32	103	102	104
19	104	116	103
2	107	107	107
35	107	122	105
29	109	107	109
23	110	109	110
34	111	112	110
21	112	115	110
33	113	133	111
9	113	112	114
8	113	116	112
27	114	118	110
15	114	131	112
37	114	127	107
1	114	121	112
38	114	121	112
36	115	113	115
12	116	122	108
22	120	115	122
7	120	121	120
39	122	113	123
30	124	127	122
14	128	127	129
24	129	126	134
16	130	133	128
26	139	137	140
17	145	142	147
25	145	147	141
31	147	168	139
20	148	151	146
5	149	179	148
Mediana	114	117	111
Média	116	120	115

As velocidades de leitura têm distribuição normal, o que foi confirmado, estatisticamente, pelo teste de Shapiro-Wilk e, visualmente, pelo histograma e diagrama de caixa da **Figura 14**.

Figura 14 – Distribuição das taxas de elocução (palavras/minuto)



O teste t de Student não revela diferença significativa entre as médias de taxa de elocução das mulheres (123,8) e homens (113,8), com um *p-value* de 0,19.

O teste T pareado comparando as velocidades de leitura nos períodos I e C confirmam que a velocidade é um pouco maior nos períodos C (*p-value* = 0.00342, intervalo de confiança da diferença de velocidade: 1,8 a 8,5 palavras/minuto). Como a taxa de elocução foi calculada contando as palavras escritas e o tempo total de duração dos períodos (palavras escritas/tempo), há dois fatores que devem ter reduzido a velocidade calculada para os períodos I:

- a) O numerador **não inclui** eventuais retratações (quando o leitor percebe a incongruência e tenta corrigi-la, repetindo ou acrescentando palavras na sua fala)
- b) O denominador **inclui** o tempo de duração de pausas e hesitações

Ou seja, as taxas de elocução dos períodos I (desconsiderando o tempo necessário para retratações, pausas e hesitações) devem ser, na realidade, maiores do que os valores que calculamos, o que nos desautoriza a considerar como significativa a diferença observada entre as taxas de elocução (taxa de elocução maior nos períodos C). Além disso, não parece fazer sentido acreditar que a taxa de elocução de cada falante varie muito ao longo dos discursos

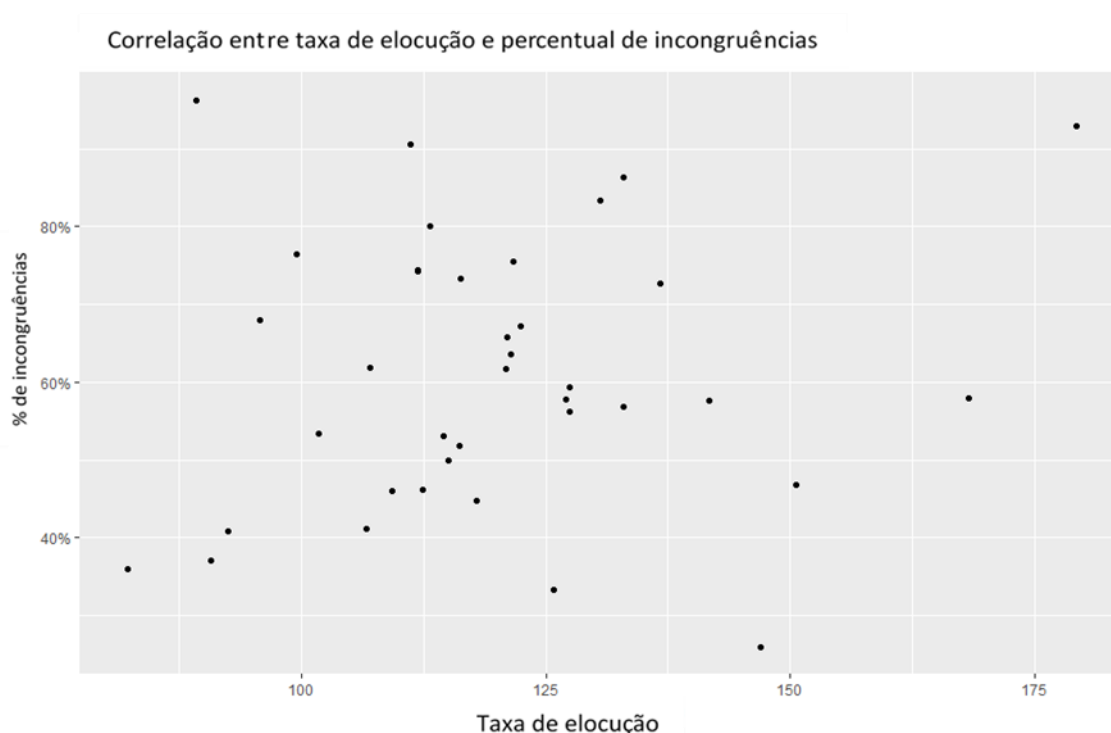
Como é metodologicamente complicado comparar a taxa de elocução entre períodos congruentes e incongruentes, podemos, alternativamente, testar a hipótese de que falantes que optam por falar mais rápido estão mais sujeitos a produzir incongruências. Essa hipótese pode ser testada confrontando a taxa de elocução média de cada falante (variável independente) com o percentual de períodos com incongruência (variável dependente). Optamos por utilizar a taxa de elocução sem incongruências como indicador da velocidade habitual de leitura do falante, pois, como já comentamos, ela reflete melhor a velocidade real de leitura. Esses dados estão descritos na **Tabela 9**.

Tabela 9 - Percentual de incongruências vs. taxa de elocução, ordenado segundo a taxa de elocução

Falante	% incongruências	Taxa de elocução
5	93%	179
31	58%	168
20	47%	151
25	26%	147
17	58%	142
26	73%	137
33	86%	133
16	57%	133
15	83%	131
30	59%	127
37	56%	127
14	58%	127
24	33%	126
35	67%	122
6	76%	122
8	64%	121
1	66%	121
38	62%	121
27	45%	118
19	73%	116
9	52%	116
22	50%	115
21	53%	115
39	80%	113
34	46%	112
36	74%	112
10	74%	112
18	91%	111
23	46%	109
29	62%	107
2	41%	107
32	53%	102
3	76%	99
7	68%	96
11	41%	93
12	37%	91
28	96%	89
13	36%	82

Para observar melhor esses dados, elaboramos um gráfico de correlação entre taxa de elocução e percentual de períodos com incongruência (**Figura 15**).

Figura 15 – Correlação entre taxa de elocução (palavras por minuto) e percentual de incongruências (número de períodos incongruentes/número total de períodos lidos)



O que observamos no gráfico da **Figura 15** é um padrão bastante disperso, que não sugere correlação significativa entre taxa de elocução e ocorrência de incongruências. Apesar disso, resolvemos testar numericamente a hipótese. Havendo testado a normalidade da distribuição do percentual de incongruências (por meio de histograma e do teste de Shapiro-Wilk), aplicamos aos dados da **Figura 15/Tabela 9** o teste de Pearson para correlação entre amostras pareadas e obtivemos um coeficiente de correlação positivo de 0,125, sugerindo uma correlação positiva fraca entre velocidade e erros, mas com um intervalo de confiança entre -0,15 e 1, e um *p-value* de 0,2273. Portanto, com os dados de que dispomos, realmente não é possível afirmar que haja correlação entre a taxa de elocução dos falantes e o percentual de incongruências, refutando nossa hipótese de que haveria mais incongruências em falas lidas com taxa de elocução mais alta.

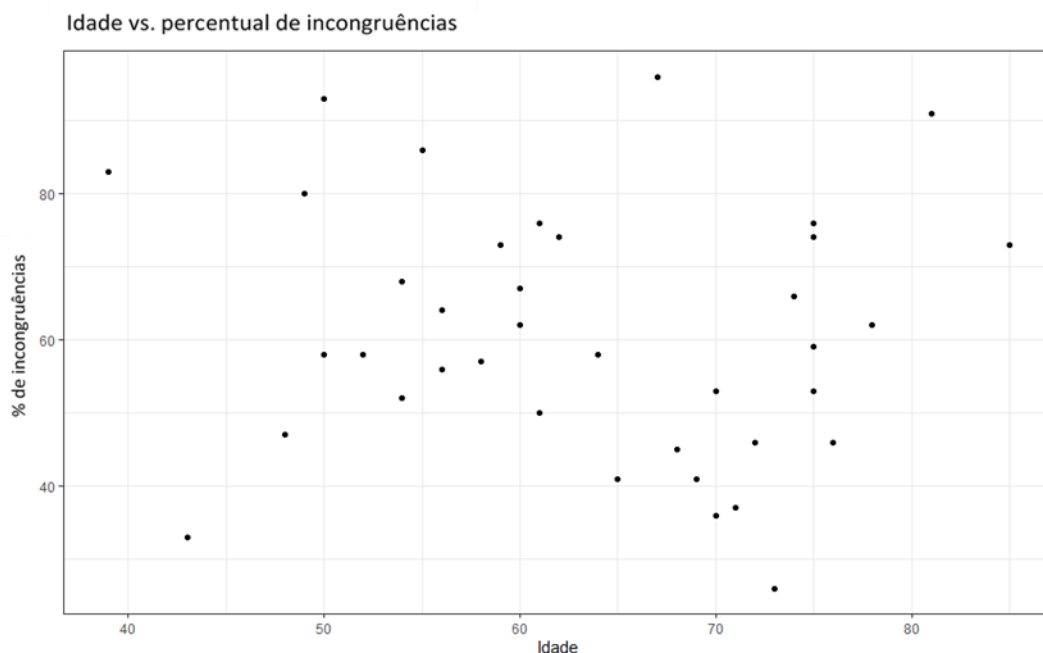
4.1.3.4 Congruência vs. idade

Outra hipótese que decidimos testar é a de que possa haver correlação entre idade (variável independente) e percentual de incongruências. A hipótese seria a de que falantes mais idosos pudessem produzir maior quantidade de incongruências durante a leitura. A **Tabela 10** descreve os dados e a **Figura 16** mostra seu gráfico de correlação.

Tabela 10 - Idade versus percentual de incongruências, ordenado pelo percentual de incongruências

Falante	Idade	Percentual de incongruências
28	67	96%
5	50	93%
18	81	91%
33	55	86%
15	39	83%
39	49	80%
3	75	76%
6	61	76%
10	75	74%
36	62	74%
19	85	73%
26	59	73%
7	54	68%
35	60	67%
1	74	66%
8	56	64%
29	78	62%
38	60	62%
30	75	59%
31	52	58%
14	50	58%
17	64	58%
16	58	57%
37	56	56%
32	70	53%
21	75	53%
9	54	52%
22	61	50%
20	48	47%
34	76	46%
23	72	46%
27	68	45%
2	69	41%
11	65	41%
12	71	37%
13	70	36%
24	43	33%
25	73	26%

Figura 16 – Idade vs. Percentual de incongruências



O gráfico da **Figura 16** não indica correlação entre idade e incongruências, e o teste de Pearson para amostras pareadas também não sugere tal correlação (p -value: 0.7241).

4.1.3.5 Congruência vs. sexo

Para completar as possíveis correlações entre dados demográficos e incongruências, não podíamos ignorar a possível correlação entre sexo e percentual de incongruências. De todos os períodos produzidos pelas mulheres, 46% continham incongruências. Quanto aos homens, 65% dos períodos que eles produziram possuíam incongruências. O teste t modificado de Welch para duas amostras revela que a diferença entre a média de períodos incongruentes produzidos pelas mulheres (46%) e a média de períodos incongruentes produzidos pelos homens (65%) é significativa (p -value = 0,0041). Ou seja, as mulheres de nossa amostra produziram proporcionalmente menos períodos incongruentes do que os homens.

4.1.4 Os padrões respiratórios

Analisamos os padrões respiratórios dos falantes de acordo com o procedimento explicado na metodologia. A partir desse método, acreditamos ter conseguido atingir um nível de detecção de

inspirações e expirações bastante próximo da realidade. Apesar disso, optamos por excluir 5 falantes (3, 4, 11, 14 e 34) cujas análises levantaram certa dúvida quanto à ocorrência de alguns movimentos respiratórios.

Os resultados gerais estão descritos na **Tabela 11**, cujas colunas indicam o seguinte:

- Tempo entre respirações: tempo total de fala dividido pelo número total de respirações; inclui, pelas características do método, o tempo respiratório.
- Palavras entre respirações: o número de palavras faladas entre duas inspirações adjacentes
- Taxa de elocução: em palavras por minuto
- Identificação dos falantes, idade e sexo

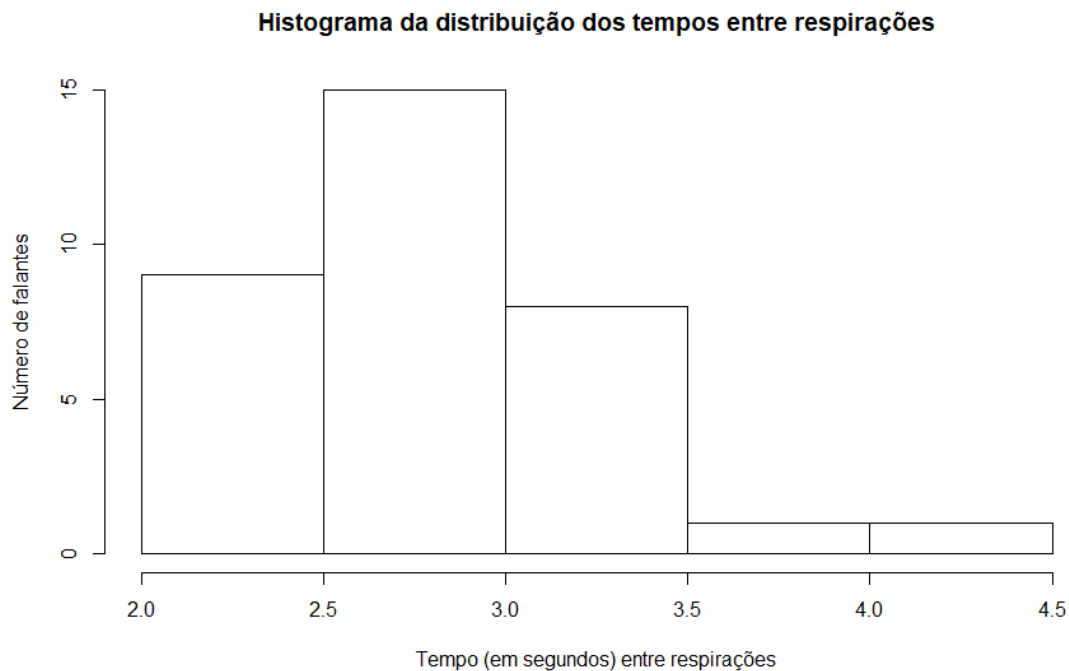
Tabela 11 - Padrões respiratórios por falante

Falante	Tempo entre respirações (s)	Palavras entre respirações	Taxa de elocução	Idade	Sexo
1	3.29	6.27	114	74	M
2	2.91	5.18	107	69	M
5	2.36	5.85	149	50	M
6	3.14	5.15	99	61	M
7	2.44	4.88	120	54	M
8	2.91	5.50	113	56	M
9	2.58	4.86	113	54	M
10	2.55	3.69	87	75	M
12	3.60	6.96	116	71	M
13	2.53	3.54	84	70	M
15	2.68	5.08	114	39	M
16	2.81	6.07	130	58	M
17	3.32	8.01	145	64	F
18	2.67	4.29	96	81	M
19	2.56	4.44	104	85	M
20	3.36	8.27	148	48	M
21	2.17	4.05	112	75	M
22	2.13	4.26	120	61	F
23	3.02	5.53	110	72	F
24	2.37	5.11	129	43	F
25	4.20	10.15	145	73	F
26	2.91	6.76	139	59	M
27	2.46	4.66	114	68	M
28	3.11	4.94	95	67	M
29	2.02	3.67	109	78	M
30	3.47	7.15	124	75	M
31	3.12	7.66	147	52	M
32	2.38	4.09	103	70	F
33	2.78	5.23	113	55	M
35	2.08	3.73	107	60	M
36	2.84	5.43	115	62	M
37	2.74	5.22	114	56	F
38	2.90	5.54	114	60	M
39	2.61	5.31	122	49	M
Média	2.79	5.21*			
Desvio padrão	0.47	0.95*			
Mediana	2.76	5.20			

*excluídos os 5 falantes com as maiores taxas de elocução e os 5 com as menores taxas de elocução, para obter uma amostra com distribuição normal

A variável “tempo entre respirações” tem distribuição normal, confirmada pelo teste de Shapiro-Wilk ($p\text{-value} = 0,3179$). A variável “palavras entre respirações”, por outro lado, têm distribuição não normal ($p\text{-value} = 0,006388$). Para poder utilizar testes estatísticos paramétricos com essa variável, excluimos os falantes com as 5 maiores e as 5 menores taxas de elocução, obtendo uma amostra de 24 falantes cuja variável “palavras entre respirações” revelou-se normal (Shapiro-Wilk com $p\text{-value} = 0,4129$). A **Figura 17** e a **Figura 18** mostram as distribuições do “tempo entre respirações” e das “palavras entre respirações”, incluindo a amostra completa, com 34 falantes.

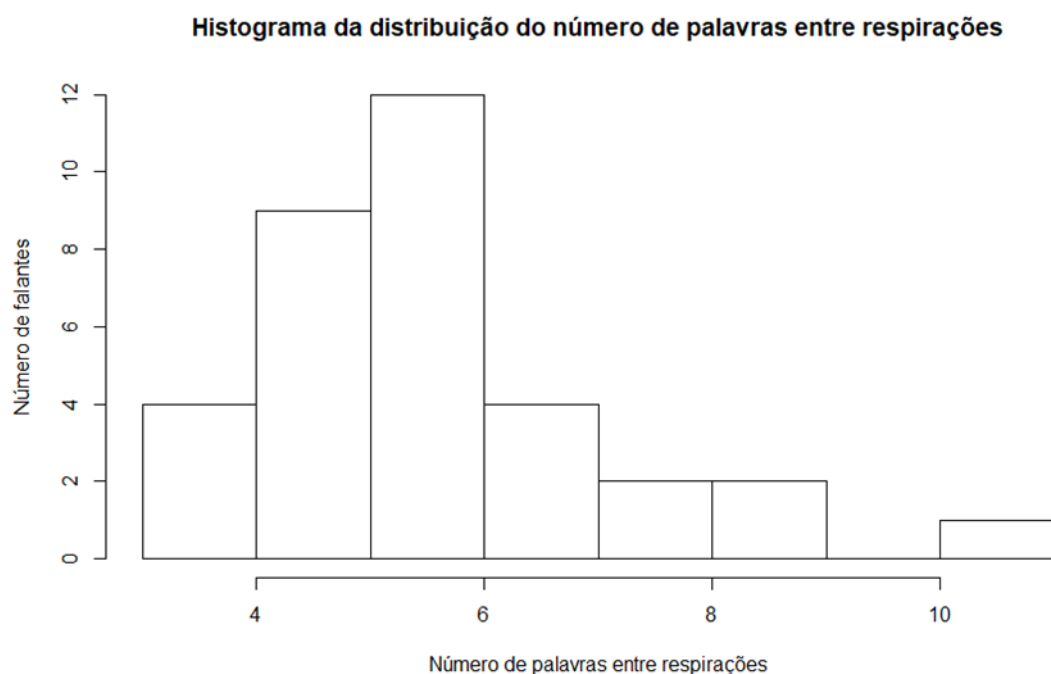
Figura 17 – Tempos entre respirações



A não normalidade da distribuição do número de palavras entre respirações deve estar relacionada ao fato de termos reunido, por questões metodológicas práticas, os dados tanto de períodos congruentes, quanto incongruentes. E, de fato, uma análise exclusiva dos períodos congruentes, considerados globalmente (sem segregação por falante), revela um **tempo entre respirações** de 3,33 segundos e de **palavras entre respirações** de 6,54 palavras. Essa diferença pode estar relacionada, entre outras coisas, às incongruências que acrescentam tempo de fala ao período, como pausas, hesitações, e todos os tipos de retratação (cujos acréscimos não foram computados como palavras faladas). Além disso, observamos que há um falante, um outlier no extremo direito da **Figura 18**, com cerca de 10 palavras entre respirações, que deve ter contribuído para a não normalidade da amostra.

Interessante notar que esse outlier é um indivíduo do sexo feminino e que foi considerado, por nossos critérios, como o participante mais desenvolvido e mais congruente entre todos os 39 falantes.

Figura 18 – Número de palavras entre respirações



Feitas essas ressalvas, realizamos as seguintes análises:

- Correlação entre **idade** e **tempo entre respirações**:
 - Resultado: não há correlação
 - Coeficiente de correlação de Pearson entre -0.25 e 0.42, *p-value* = 0.5754
- Correlação entre **sexo** e **tempo entre respirações**:
 - Resultado: não há correção
 - Teste t de Student com intervalo de confiança (95%) entre -0.77 e 0.56, e *p-value* = 0.7144
- Correlação entre **taxa de elocução** e **palavras entre respirações (Figura 19)**:
 - Resultado: há correlação positiva, indicando que falantes que leem mais rápido, inserem mais palavras entre duas respirações sucessivas
 - Coeficiente de correlação de Pearson = 0,57, *p-value* = 0,003

- Correlação entre **taxa de elocução** e **tempo entre respirações** (Figura 20):
 - Resultado: não há correlação, indicando que, independentemente da taxa de elocução, os falantes mantêm seu padrão respiratório
 - Coeficiente de correlação de Pearson = 0,21, *p-value* = 0,33
- Correlação entre **tempo entre respirações** e **palavras entre respirações** (Figura 21)
 - Resultado: há correlação positiva entre as variáveis, indicando que os falantes que respiram menos vezes, inserem mais palavras entre duas respirações sucessivas.
 - Coeficiente de correlação de Pearson = 0,92 e *p-value* = 0,00000000024

Figura 19 – Correlação entre “taxa de elocução” e “número de palavras entre respirações”

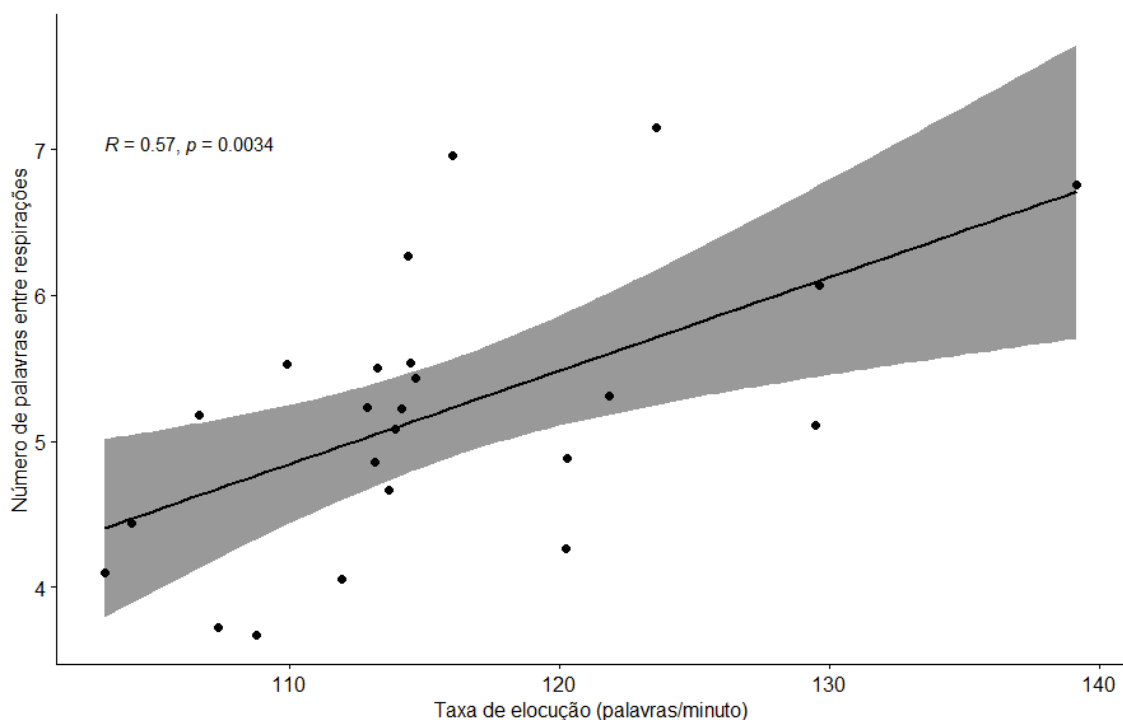


Figura 20 – Correlação entre “taxa de elocução” e “tempo entre respirações”

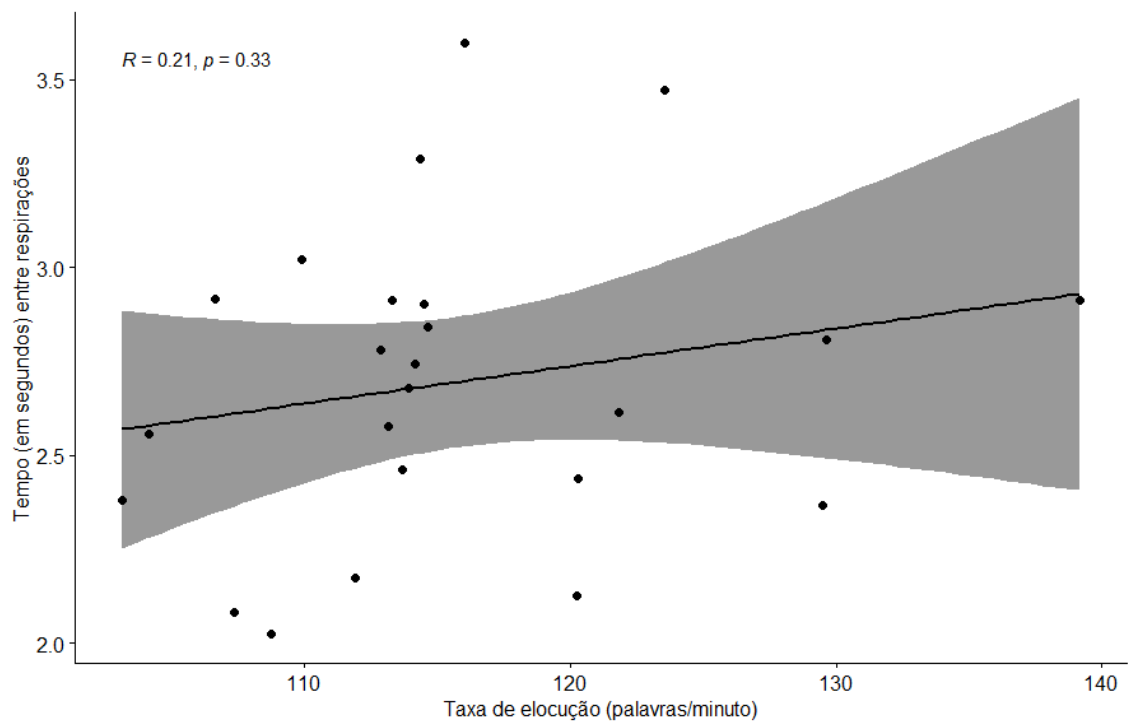
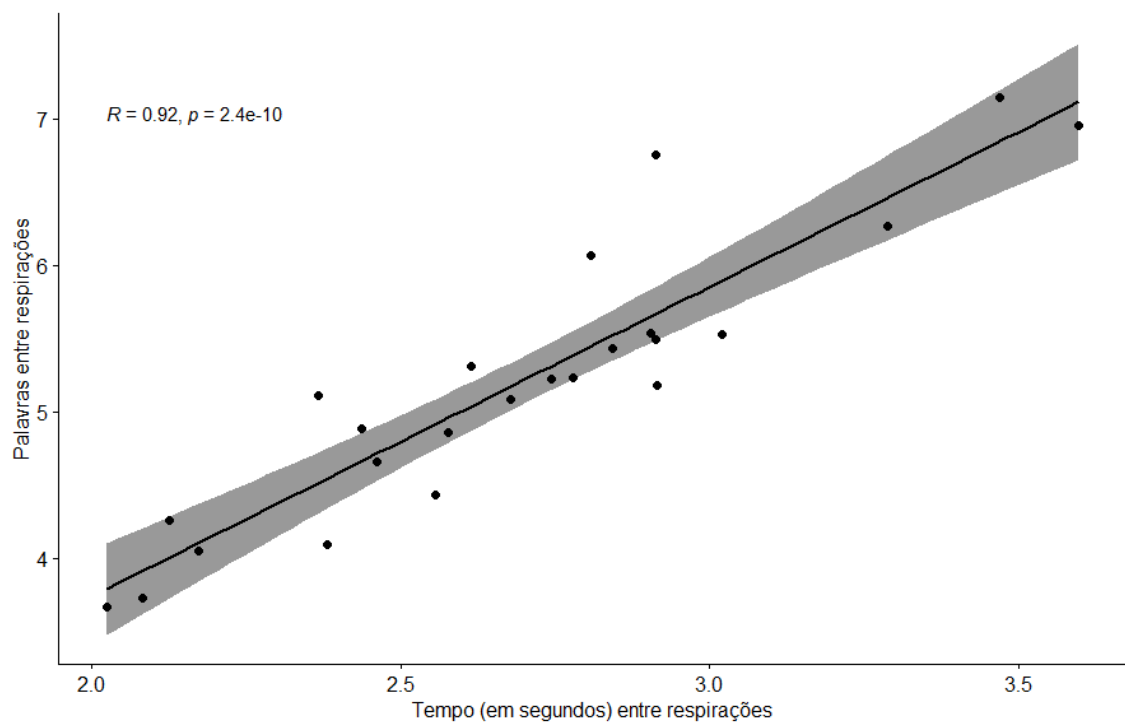


Figura 21 - Correlação entre tempo entre respirações e palavras entre respirações



Além do que descrevemos e analisamos, achamos que seria útil acrescentar aos resultados um fenômeno adicional que observamos durante o estudo das respirações. Percebemos que há respirações completas, com ampla inspiração, e respirações incompletas, muito rápidas, com volume inspiratório mínimo, quase irrisório. Essas últimas sugerem um tipo “apressado” de inspiração, pois costumam aparecer imediatamente antes de alguma emissão sonora. Por exemplo, no período abaixo, marcamos as respirações normais (R) e a respiração “apressada” (r):

- (R)Um retorno ao parlamento, nesta altura da vida, (R)traz à tona certas reminiscências, (R)certas recordações de outros tempos, (R)às quais (r)é difícil resistir.

Deve-se acrescentar que a (r) apareceu logo depois de uma incongruência (troca de “às quais” por “aos quais”), mas em numerosos eventos de (r), não houve incongruência associada. Infelizmente, por questões de tempo e de prazo, não foi possível elaborar um estudo mais detalhado desse fenômeno, que talvez pudesse ser classificado como ‘pausa silenciosa’ (R) e ‘pausa respiratória’ (r).

4.1.5 Resumo da seção 4.1

Em resumo, os resultados descritos nesta seção 4.1, que trata dos falantes, do corpus, das incongruências *lato sensu* e dos padrões respiratórios mostram que falantes de PB, predominantemente homens, com nível de escolaridade majoritariamente superior, nascidos em diversos estados brasileiros, com idade mediana de 62 anos, que empreenderam leituras públicas com duração mediana de 8 minutos, produziram incongruências em mais da metade dos períodos lidos.

A idade dos falantes e a taxa de elocução não influenciaram o percentual de períodos incongruentes, mas o sexo e a quantidade de palavras em cada período, sim. As mulheres produziram menos incongruências que os homens. E, independentemente do sexo, quanto mais longos são os períodos, mais incongruências são encontradas. Observou-se, também, que a relação entre o número de palavras e de incongruências é linear; ou seja, cada palavra a mais acrescenta uma probabilidade semelhante de ocorrência incongruências (tanto segmentares, quanto suprasegmentares).

E o estudo dos padrões respiratórios sugere que a variabilidade das taxas de elocução entre os falantes se dá pela inserção de maior número de palavras entre cada respiração, não havendo aumento da frequência respiratória dos falantes mais “velozes”. Além disso, nem a idade, nem o sexo influenciaram a frequência respiratória.

4.1.6 Prólogo das próximas seções (4.2 e 4.3)

Na sequência, abordaremos, sucessivamente, os 1363 eventos de incongruências segmentares (seção 4.2) e os 858 eventos de incongruências suprasegmentares (seção 4.3) identificados no corpus. Uma visão geral dessas incongruências está ilustrada na **Tabela 12** e na **Tabela 13**. A primeira **tabela** mostra as incongruências ordenadas por número de ocorrências. Na segunda, elas estão ordenadas por tipo: segmentar ou suprasegmentar.

Tabela 12 - Todas as incongruências do corpus

Incongruências	Eventos	Classificação
Troca de palavra	360	segmentar
Ortoépia	334	segmentar
Concordância	307	segmentar
Pausa	294	suprasegmentar
Inserção de fronteira não terminal	240	suprasegmentar
Acréscimo de palavra	172	segmentar
Omissão de palavra	146	segmentar
Inserção de fronteira terminal	117	suprasegmentar
Deleção de fronteira não terminal	70	suprasegmentar
Hesitação	51	suprasegmentar
Outras (suprasegmentares)	45	suprasegmentar
Troca de frase	21	segmentar
Deleção de fronteira terminal	21	suprasegmentar
Lentificação	17	suprasegmentar
Repetição de palavra	13	segmentar
Repetição de sílaba	10	segmentar
Afirmção lida como pergunta	2	suprasegmentar
Pergunta lida como afirmação	1	suprasegmentar
Total		2221

Tabela 13 - Todas as incongruências do corpus, ordenadas por classificação

Incongruências	Eventos	Classificação
Troca de palavra	360	segmentar
Ortoépia	334	segmentar
Concordância	307	segmentar
Acréscimo de palavra	172	segmentar
Omissão de palavra	146	segmentar
Troca de frase	21	segmentar
Repetição de palavra	13	segmentar
Repetição de sílaba	10	segmentar
Total		1363
Pausa	294	suprasegmentar
Inserção de fronteira não terminal	240	suprasegmentar
Inserção de fronteira terminal	117	suprasegmentar
Deleção de fronteira não terminal	70	suprasegmentar
Hesitação	51	suprasegmentar
Outras (suprasegmentares)	45	suprasegmentar
Deleção de fronteira terminal	21	suprasegmentar
Lentificação	17	suprasegmentar
Afirmção lida como pergunta	2	suprasegmentar
Pergunta lida como afirmação	1	suprasegmentar
Total		858

4.2 Descrições e análises das incongruências segmentares

As descrições e análises das incongruências segmentares organizam-se da seguinte forma:

- 4.2.1 As incongruências segmentares
- 4.2.2 As retratações
- 4.2.3 As trocas de palavras
 - 4.2.3.1 A frequência das palavras nos corpora
 - 4.2.3.2 As categorias gramaticais
 - 4.2.3.3 As semelhanças entre palavras originais e inseridas
- 4.2.4 Resumo da seção

4.2.1 As incongruências segmentares

Nesta seção, descrevemos e analisamos os 1363 eventos de incongruências segmentares. Esses eventos, distribuídos por categoria e ordenados pelo número de ocorrências, estão discriminados na **Tabela 14** - Incongruências segmentares.

Tabela 14 - Incongruências segmentares

Incongruência	Ocorrências	%
Troca de palavra	360	26%
Ortoépia	334	25%
Concordância	307	23%
Acréscimo de palavra	172	13%
Omissão de palavra	146	11%
Troca de frase	21	2%
Repetição de palavra	13	1%
Repetição de sílaba	10	1%
Total	1363	100%

Com relação à categoria “concordância”, suas incongruências foram discriminadas em subcategorias de gênero e de número, conforme a discordância com o texto escrito tenha se materializado como masculinização ou feminilização de palavras, respectivamente, femininas ou masculinas; ou pluralização/singularização de palavras singulares/plurais. Consideramos todas as não concordâncias, tanto nominais, quanto verbais. A **Tabela 15** mostra a distribuição das categorias gramaticais das palavras envolvidas nas incongruências de concordância.

Tabela 15 - Categorias gramaticais nas incongruências de concordância

Categoria gramatical	Eventos
substantivo	103
adjetivo	57
particípio	37
verbo	29
preposição + artigo definido	25
artigo definido	16
outros	39

A **Tabela 16** mostra o número de eventos por subcategoria, onde é possível notar:

- a predominância da discordância de número sobre a discordância de gênero
- a predominância de singularização (de palavras que, no texto original, estavam no plural) sobre a pluralização

Tabela 16 - Desvios de concordância

Tipo	Eventos	%
singularização	187	60.9%
pluralização	77	25.1%
masculinização	21	6.8%
feminilização	18	5.9%
fem+plur	2	0.7%
masc+plur	1	0.3%
masc+sing	1	0.3%

fem: feminilização; masc: masculinização; sing: singularização; plur: pluralização

Além de identificar as incongruências segmentares já descritas, identificamos também dois fenômenos adicionais que se associaram a essas incongruências. O primeiro é que, em determinados contextos, a incongruência não causa prejuízo significativo ao período, como exemplificado abaixo e descrito, a seguir, na **Tabela 17**.

Tabela 17 - Incongruências segmentares **com** e **sem** prejuízo ao período

Incongruências	Total	Com prejuízo	%	Sem prejuízo	%
Ortoépia	334	334	100%	0	0%
Troca de palavra	360	300	83%	60	17%
Concordância	307	280	91%	27	9%
Omissão de palavra	146	113	77%	33	23%
Acréscimo de palavra	172	89	52%	83	48%
Troca de frase	21	18	86%	3	14%
Repetição de palavra	13	13	100%	0	0%
Repetição de sílaba	10	10	100%	0	0%
Total	1363	1157	-	206	-

Nesses casos, a leitura realizada é gramatical e aceitável, mas nós não sabemos se o leitor conseguiu fazer essa avaliação durante a leitura; se ele fez, a alteração foi voluntária; se ele não fez, o

período produzido estaria “acidentalmente” correto. A seguir, descrevemos algumas situações concretas ilustrativas desse raciocínio:

- Concordância:
 - Texto escrito: “...são **registradas** também agressões físicas, abusos, maus tratos, negligência e abandono.”
 - Fala lida: “são **registrados** também agressões físicas, abusos, maus tratos, negligência e abandono.
- Acréscimo de palavra:
 - Texto escrito: “Ele segue tentando pagar **as** dívidas, ...”
 - Fala lida: “Ele segue tentando pagar **suas** dívidas, ...”
- Omissão de palavra:
 - Texto escrito: “Se, por alguma razão **qualquer**, o Brasil fosse atingido...”
 - Fala lida: “Se, por alguma razão, o Brasil fosse atingido...”
- Troca de palavra:
 - Textos escrito: “...que os princípios inscritos na Constituição Federal de 1988 e legislação posterior sejam **de fato** respeitados.
 - Fala lida: “...que os princípios inscritos na Constituição Federal de 1988 e legislação posterior sejam **efetivamente** respeitados.”

Na **Tabela 17**, é possível observar que, das 1363 incongruências segmentares observadas, 1157 foram prejudiciais aos períodos. Se considerarmos apenas as incongruências prejudiciais, os desvios de “ortoépia” são as mais frequentes, superando a “troca de palavras”. Por definição, as incongruências de “ortoépia” (bem como a repetição de sílabas) sempre causam prejuízo (100% de prejuízo).

4.2.2 As retratações

O segundo fenômeno adicional que resolvemos registrar, conforme dissemos há pouco, é que, muitas vezes, quando há incongruência **com** prejuízo ao período, o falante se retrata, retificando a incongruência. Ou seja, ele produz a incongruência, percebe que a cometeu e, na sequência, a corrige. A **Tabela 18** mostra as retratações ocorridas em cada categoria de incongruência segmentar.

Tabela 18 – Retratações em incongruências segmentares **com** prejuízo ao período

Incongruências	Com prejuízo	Retratações	%
Ortoépia	334	163	49%
Troca de palavra	300	97	32%
Concordância	280	26	9%
Omissão de palavra	113	13	12%
Acréscimo de palavra	89	2	2%
Troca de frase	18	4	22%
Repetição de palavra	13	0	0%
Repetição de sílaba	10	0	0%
Total	1157	305	-
%	100%	26%	

Nesta **Tabela 18**, observamos que os desvios de ortoépia são corrigidos em metade dos casos, contra apenas um terço dos casos nas trocas de palavras, 9% nos desvios de concordância e 2% nos acréscimos de palavra. Em apenas três casos (não descritos na **Tabela 18**), houve retratação **sem** ter havido prejuízo ao período, todos em situação de troca de palavra. O teste do qui-quadrado aplicado à **Tabela 18** (excluídas as três últimas linhas, que tiveram relativamente poucas ocorrências) confirma que a incidência de retratações é diferente entre as diversas categorias de incongruências segmentares ($p\text{-value} < 2.2e-16$). Para identificar quais discrepâncias foram responsáveis pela diferença, analisamos os resíduos e os $p\text{-values}$ em uma análise post-hoc e observamos que há proporcionalmente mais retratações nos desvios de ortoépia do que nas demais categorias, e proporcionalmente menos retratações nos desvios de concordância, e nas omissões e acréscimos de palavras. No caso das trocas de palavras, a incidência de retratações está na faixa intermediária entre as demais categorias e foi considerada pelo teste como não significativa.

Se retirarmos do cômputo das incongruências segmentares aquelas que foram retificadas (retratação), restam-nos as incongruências cujos prejuízos permaneceram na fala lida; ou seja, os casos em que houve prejuízo e não houve retratação. A **Tabela 19** descreve tais incongruências.

Tabela 19 - Incongruências segmentares **com** prejuízo e **não** retratadas

Incongruências	Total	Não retratadas	%
Ortoépia	334	171	51%
Troca de palavra	300	203	68%
Concordância	280	254	91%
Omissão de palavra	113	100	88%
Acréscimo de palavra	89	87	98%
Troca de frase	18	14	78%
Repetição de palavra	13	13	100%
Repetição de sílaba	10	10	100%
Total	1157	852	-

Seria possível, também, agrupar as categorias de incongruências segmentares que descrevemos acima em duas classes, segundo o nível em que as alterações ocorrem com relação ao texto escrito. A primeira classe, que chamaremos arbitrariamente de frasal, envolve modificações na estrutura de frase/oração/período por meio de trocas, omissões ou acréscimos de palavras inteiras. A segunda classe, que chamaremos, também arbitrariamente, de lexical, envolve modificações na estrutura interna das palavras, que são flexionadas de forma gramaticalmente inadequada (concordância), ou são pronunciadas de forma anômala (ortoépia). Esse agrupamento das categorias de incongruências segmentares está descrito na **Tabela 20**, onde observamos uma distribuição equilibrada entre esses dois tipos arbitrários de incongruência.

Tabela 20 - Incongruências segmentares frasais vs. lexicais

Incongruência	Ocorrências	Incongruência	Ocorrências	%
Troca de palavra	360			
Acréscimo de palavra	172			
Omissão de palavra	146	Frasal	709	52%
Troca de frase	21			
Repetição de palavra	10			
Ortoépia	334			
Concordância	307	Lexical	654	48%
Repetição de sílaba	13			
Total	1363		1363	100%

Um fenômeno que observamos ao analisar as retratações foi que, em muitos casos, o falante, quando se retrata, inclui palavras anteriores à palavra incongruente em sua retratação. Por exemplo, no trecho “...geração de caixa **no curto** prazo...”, o falante disse “...geração de caixa **no cum...no curto** prazo...”. Ele foi incongruente na leitura da palavra “**curto**”, percebeu a incongruência, e se retratou;

mas não retratou apenas a palavra incongruente: incluiu na retratação uma palavra congruente anterior à incongruência (“no”).

A seguir, listamos alguns exemplos adicionais do fenômeno de retratação com adição de palavras não incongruentes. Retiramos do corpus alguns trechos dos textos escritos e suas leituras correspondentes. Destacamos em negrito as palavras que foram incluídas na retratação (repetidas), apesar de não terem sido lidas com incongruência:

- Texto
 - a) Os saudosos fundadores...
 - b) A presença ilustre e festejada...
 - c) ...a defesa do consumidor como um dos princípios...
 - d) ...afiliadas e retransmissoras...
 - e) ...para que o consórcio conclua...
 - f) ...cada país deve lidar com...
- Leitura
 - a) Os sauda...**os** saudosos fundadores...
 - b) A presença ilustre e festeje...**e** festejada...
 - c) c
 - d) ...afiliadas e ran...**e** retransmissoras...
 - e) ...para que o concór...**para que o** consórcio conclua...
 - f) ...cada país deve vilili...**cada país deve** lidar com...

A **Tabela 21** mostra a ocorrência desse fenômeno (adição de palavras não incongruentes à retratação de uma palavra incongruente) entre as incongruências segmentares mais comuns (ortoépia e troca de palavra).

Tabela 21 - Retratações com ou sem adição de palavras não incongruentes

Adição de palavra(s)	Sim	%	Não	%
Troca de palavra	31	34%	63	39%
Ortoépia	60	66%	98	61%
%	36%		64%	

Esses dados mostram que o fenômeno é relativamente comum, ocorrendo em 36% das retratações analisadas. Além disso, sua incidência não é influenciada pelo tipo de incongruência, uma vez que ela ocorre com frequência semelhante tanto nas trocas de palavra, quanto nas incongruências

por desvio de ortoépia (*p-value* de 0,5074 no *Pearson's Chi-squared test with Yates' continuity correction*).

Como se trata de um fenômeno intrigante, resolvemos aprofundar sua investigação. Com base em uma análise preliminar, pareceu-nos que nos casos em que havia adição de palavras à retratação havia também predominância de certas categorias gramaticais e alguma diferença no tamanho das palavras. Identificamos, então (nas incongruências caracterizadas por troca de palavra ou desvio de ortoépia com ocorrência de retratação), as categorias gramaticais e o tamanho (em sílabas) das seguintes palavras envolvidas no fenômeno de retratação:

- A palavra incongruente retratada
- As três palavras congruentes imediatamente anteriores à palavra incongruente retratada

Com relação à categoria gramatical das palavras incongruentes retratadas, sua distribuição está ilustrada na **Tabela 22**.

Tabela 22 - Categoria gramatical das palavras incongruentes retratadas, com e sem adição de palavra(s) congruentes(s) à retratação

Categoria	Com adição	%	Sem adição	%
substantivo	56	57.1%	42	26.1%
adjetivo	15	15.3%	32	19.9%
verbo	11	11.2%	32	19.9%
numeral	6	6.1%	13	8.1%
outros	5	5.1%	4	2.5%
art. definido (ad)	1	1.0%	2	1.2%
advérbio	1	1.0%	8	5.0%
conjunção	1	1.0%	2	1.2%
preposição (prep)	1	1.0%	13	8.1%
prep+ad	1	1.0%	13	8.1%

Observam-se nessa **Tabela 22** algumas discrepâncias nas distribuições das categorias gramaticais das palavras incongruentes retratadas quando houve adição e quando não houve adição de outras palavras à retratação:

- Predomínio de substantivos em “com adição”
- Incidência relativamente alta de preposições em “sem adição”
- Uma possível desproporção na incidência de adjetivos e verbos em “sem adição”

De fato, o teste do qui-quadrado confirma que as distribuições não são iguais (Pearson's Chi-squared test with simulated p -value = 0,0004998). Para identificar melhor onde estão essas diferenças estatísticas, agrupamos os dados da **Tabela 22** em apenas quatro categorias (substantivo, verbo, adjetivo e outras), incluindo na categoria “outras” todos os dados não classificados como verbo, adjetivo ou substantivo. Com isso, eliminamos as categorias gramaticais com poucas ocorrências, para poder realizar o teste do qui-quadrado sem a necessidade de simular o p -value. Os dados agrupados estão na **Tabela 23**.

Tabela 23 - Categoria gramatical da palavra retratada com ou sem adição de palavra(s) congruentes(s) à retratação
Agrupamento em quatro categorias

Categoria	Com adição	%	Sem adição	%
substantivo	56	57.1%	42	26.1%
adjetivo	15	15.3%	32	19.9%
verbo	11	11.2%	32	19.9%
outros	16	16,2%	55	34,2%

Com esse agrupamento, obtivemos um p -value de 0,006355 (Pearson's Chi-squared test), reforçando a ideia de que existe uma associação entre a categoria gramatical da palavra retratada e a adição de palavras à retratação.

O próximo passo foi analisar a distribuição de categorias gramaticais das palavras congruentes que foram adicionadas à retratação da palavra incongruente. Antes disso, entretanto, devemos mencionar dois fenômenos que notamos e que optamos por excluir da análise:

- Onde houve **vírgula** imediatamente antes da palavra incongruente retratada (13 eventos), não houve nenhuma adição de palavras congruentes à retratação; portanto, excluimos da análise este tipo de situação
- Os **numerais** (também excluídos) revelaram-se uma categoria de análise diferente das demais, devido a:
 - terem sido escritos, na maioria das vezes, com algarismos (p.ex.: 2016)
 - formarem unidades com possível coesão interna (p.ex.: dois mil e dezesseis), o que nos levou a acreditar que mereceriam uma análise mais aprofundada, não condizente com o contexto e os prazos de realização deste estudo

Outro fenômeno que observamos é que, por vezes, são adicionadas à retratação mais de uma palavra; como neste exemplo, que já mencionamos, onde houve adição de 3 palavras congruentes à

palavra incongruente: “...para que o concór... **para que o** consórcio conclua...”. A **Tabela 24** mostra a distribuição do número de palavras adicionadas à retratação.

Tabela 24 - Número de palavras adicionadas à retratação, quando houve adição de palavra(s) congruente(s) à palavra incongruente

1 palavra adicional	70 eventos
2 palavras adicionais	12 eventos
3 palavras adicionais	7 eventos

Observa-se na **Tabela 24** que, na maioria dos casos, houve adição de apenas uma palavra à palavra incongruente retratada; e que a adição de um número maior de palavras torna-se menos frequente à medida que esse número aumenta, atingindo o máximo de 3 palavras adicionais.

Considerando que a grande maioria das retratações com adição de palavra adiciona apenas uma palavra à retratação da palavra incongruente, analisamos apenas a categoria gramatical da palavra imediatamente anterior à palavra incongruente, tanto nos casos em que houve adição, como nos casos em que não houve adição de palavras à retratação. A **Tabela 25** ilustra essa distribuição.

Tabela 25 - Categoria gramatical da palavra imediatamente anterior à palavra incongruente retratada, incluindo as situações em que a palavra **foi adicionada e não foi adicionada** à retratação

Categoria gramatical	Adicionada	%	Não adicionada	%
art. def.	25	26.88%	8	5.48%
prep+ad	17	18.28%	16	10.96%
preposição	14	15.05%	22	15.07%
verbo	8	8.60%	17	11.64%
conjunção	6	6.45%	6	4.11%
art. indef.	5	5.38%	5	3.42%
"que"	4	4.30%	2	1.37%
substantivo	4	4.30%	41	28.08%
advérbio	3	3.23%	7	4.79%
outros	7	7.53%	5	3.42%
adjetivo	0	0.00%	17	11.54%

Observamos na **Tabela 25** as seguintes particularidades:

- Categorias gramaticais mais frequentes entre as palavras adicionadas, em comparação com as palavras não adicionadas:

- Artigos definidos
- Preposições, mas apenas quando em “contração” com artigo definido (p.ex.: do)
- Categorias gramaticais mais frequentes entre as palavras não adicionadas, em comparação com as palavras adicionadas:
 - Substantivos
 - Adjetivos

As discrepâncias de frequências entre as condições “com” e “sem” adição de palavras à retratação, observadas na **Tabela 25**, são estatisticamente significativas (Pearson's Chi-squared test with simulated *p-value* (based on 2000 replicates) = 0,0004998), acompanhando o mesmo padrão observado para as palavras incongruentes retratadas.

Para finalizar a análise das categorias gramaticais, classificamos as palavras incluídas nas retratações, para ver quais são as combinações de categorias gramaticais mais comumente associadas à retratação com adição de palavras. Os resultados dessa análise estão na **Tabela 26**.

Tabela 26 - Categorias gramaticais das palavras incongruentes e das palavras adicionais incluídas nas retratações

Palavra adicional (-2)	Palavra adicional (-1)	Palavra incongruente retratada	Número de ocorrências
preposição	artigo definido (ad)	substantivo	20
	prep/ad	substantivo	12
	preposição (prep)	substantivo	7
	artigo definido	substantivo	2
	artigo definido	adjetivo	2
	conjunção	adjetivo	2
	nome próprio	sobrenome	2
	part. apassivadora	verbo	2
	prep/ad	adjetivo	2
	"que"	verbo	2
	verbo	adjetivo	2
	verbo	verbo	2
	outras	(1 ocorrência cada)	32

O que vemos na **Tabela 26** é que o fenômeno de retratação de uma palavra incongruente (por troca de palavra ou desvio de ortoépia) com adição de palavras congruentes à retratação ocorre, predominantemente, quando a palavra incongruente é um substantivo e a palavra que a antecede é, em ordem decrescente de frequência, um artigo definido, uma contração preposição+artigo definido, ou uma preposição. Abaixo, fornecemos alguns exemplos concretos retirados do corpus:

- Artigo definido + substantivo:
 - ...a frinçã...a função primordial...
 - ...a prea...a produção de energia
- Preposição/artigo definido + substantivo:
 - ...saúde financeira das est...das entidades...
 - ...encaminhado à sessã...à sanção...
- Preposição + substantivo:
 - ...parcelamento de ju...de férias...
 - Sem val...sem salvaguardas...

Passamos agora, como nos propusemos acima, a analisar o tamanho (em sílabas) das palavras e a estabelecer suas eventuais correlações com o fenômeno de retratação com adição de palavras congruentes. A amostra que estudamos incluiu a palavra incongruente e a palavra imediatamente anterior. A **Tabela 27** descreve os achados, nas condições em que a palavra anterior foi adicionada à retratação, ou não foi adicionada à retratação.

Tabela 27 - Distribuição do número de palavras por tamanho (em sílabas) nas situações de retratação "com" e "sem" adição de palavras adicionais

tamanho	com	%	sem	%
Palavra incongruente				
1	7	8%	25	18%
2	27	29%	15	11%
3	18	20%	21	15%
4	25	27%	33	24%
5	9	10%	32	24%
6	4	4%	7	5%
7	2	2%	3	2%
Palavra imediatamente anterior				
1	76	84%	47	35%
2	13	14%	38	28%
3	2	2%	27	20%
4	0	0%	12	9%
5	0	0%	9	7%
6	0	0%	2	1%
7	0	0%	0	0%

Nessa **Tabela 27** é possível observar as diferenças de frequência de tamanhos de palavras (palavra retratada e palavra imediatamente anterior) nas condições “com” e “sem” palavras adicionais; e percebemos o seguinte:

- Com relação às palavras incongruentes, quando há adição da palavra anterior à retratação, parece haver uma menor incidência de palavras muito pequenas ou muito grandes, havendo predomínio de palavras médias (de 2 a 4 sílabas)
- Com relação às palavras imediatamente anteriores às palavras incongruentes, quando essas palavras são adicionadas à retratação, parece haver uma maior incidência de palavras pequenas (predominantemente, de 1 sílaba) e uma ausência absoluta de palavras com 4 sílabas ou mais

Essas diferenças nos tamanhos das palavras envolvidas nas retratações com e sem adição de palavras congruentes podem ser mais bem visualizadas nos gráficos da **Figura 22** e **Figura 23**.

Figura 22 – Número de ocorrências por tamanho da palavra incongruente

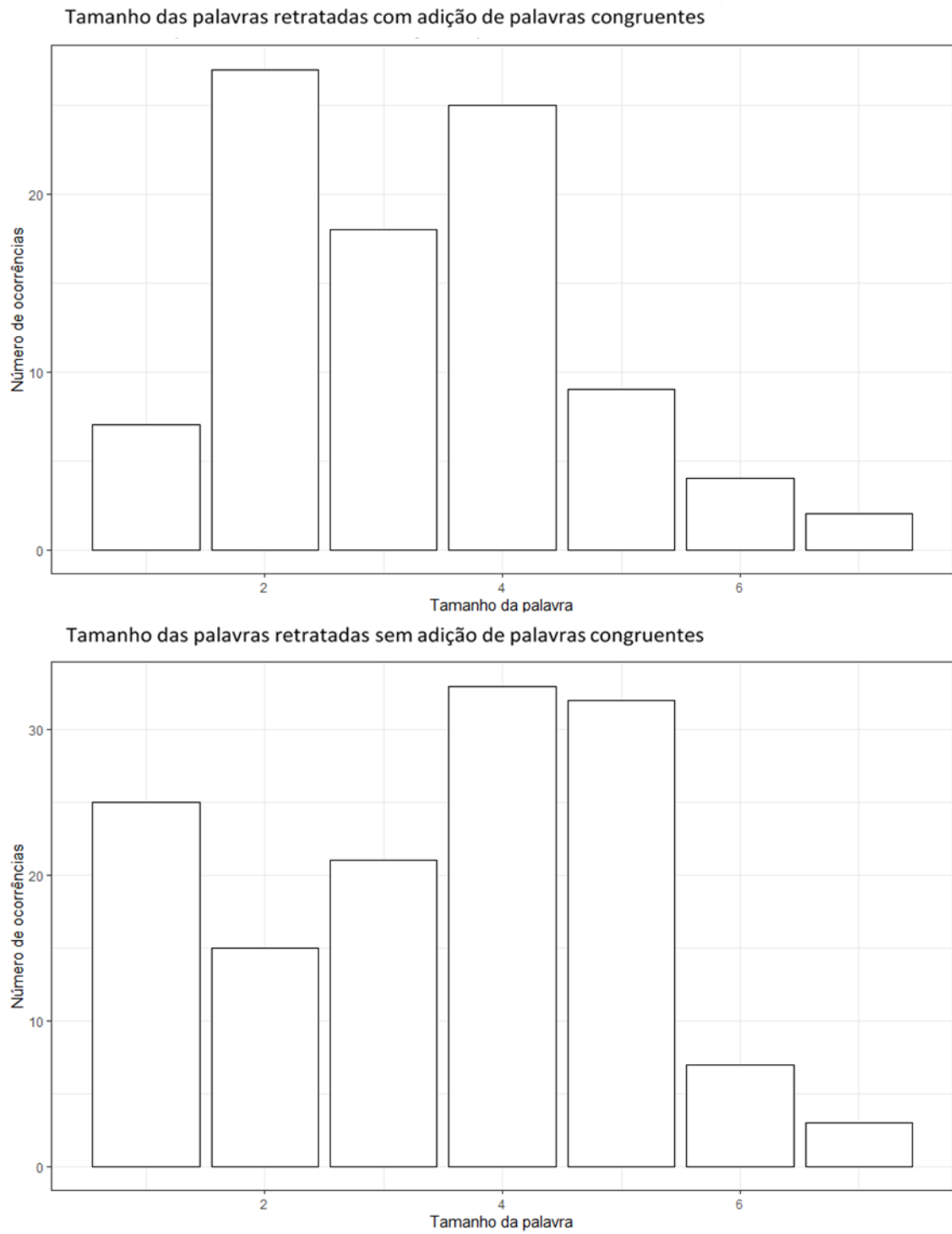
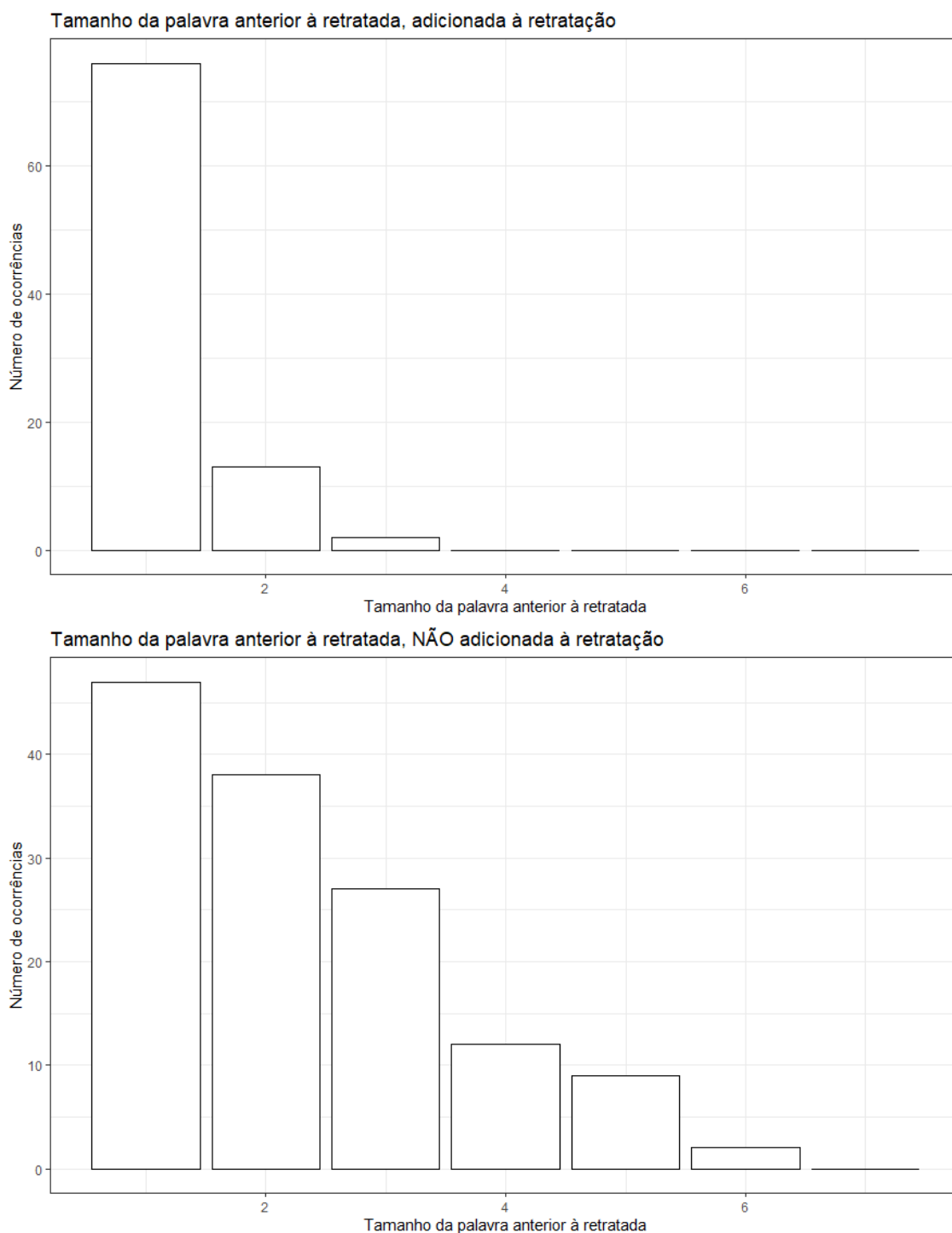


Figura 23 – Número de ocorrências por tamanho da palavra imediatamente anterior à palavra incongruente



O teste do qui-quadrado aplicado aos dados da **Tabela 27** (Pearson's Chi-squared test with simulated *p-value* (based on 2000 replicates)) apontam que as distribuições de frequência dos tamanhos das palavras nas condições com e sem adição de palavras congruentes à retratação, considerando as palavras congruentes e as palavras imediatamente anteriores, são realmente diferentes (*p-values*, respectivamente, de 0,0009995 e 0.0004998). Ou seja, os pares “palavra imediatamente

anterior” + “palavra incongruente” que são retratados em conjunto caracterizam-se por “tamanho pequeno” + “tamanho médio”, como na frase hipotética “...o chimp...o jacaré...”.

Uma dúvida que surge após essa análise inicial do tamanho das palavras é se a associação que encontramos entre os diferentes tamanhos e a maior incidência de retração com adição de palavras é realmente devida ao tamanho das palavras. Porque sabemos que substantivos, por exemplo, são, via de regra, maiores do que artigos definidos e, sendo assim, confundem-se as possíveis associações entre tamanho e categoria gramatical. Para tentar minimizar essa confusão, extraímos os resultados de apenas duas categorias gramaticais — substantivos e preposições (com e sem contração com artigo definido) — e comparamos o tamanho das palavras de cada uma dessas categorias entre si (substantivo vs. substantivo; preposição/art. def. vs. preposição/art. def.), nas situações em que houve e em que não houve adição de palavras congruentes à retração. A **Tabela 28** e a **Tabela 29** ilustram os resultados dessa análise.

Tabela 28 - Tamanho dos substantivos incongruentes nas condições "com" e "sem" adição de palavras à retração

tamanho	com		sem	
	ocorrências	%	ocorrências	%
2	16	29%	4	10%
3	13	23%	7	17%
4	16	29%	11	26%
5	6	11%	13	31%
6	3	5%	4	10%
7	2	4%	3	7%

Tabela 29 - Tamanho das preposições (com ou sem art. def.) anteriores às palavras incongruentes nas condições em que foram ou não foram adicionadas à retração

tamanho	adicionada		não adicionada	
	ocorrências	%	ocorrências	%
1	29	94%	28	74%
2	2	6%	10	26%

O teste do qui-quadrado (Pearson's Chi-squared test with simulated *p-value* (based on 2000 replicates)) sugere que as diferenças observadas são significativas (*p-value* = 0,04398 e 0,05497, respectivamente). As possíveis conclusões dessa análise são:

- substantivos menores (com até 3 sílabas) têm maior chance de serem retratados com adição da palavra anterior, quando comparados a substantivos maiores (5 ou mais sílabas)
- preposições (com ou sem contração com artigo definido) com 2 sílabas têm menor chance de serem incluídos na retratação de uma palavra incongruente subsequente do que aquelas com apenas 1 sílaba.

4.2.3 As trocas de palavra

4.2.3.1 A frequência das palavras nos corpora

A análise seguinte diz respeito às trocas de palavra. Uma hipótese que levantamos foi a de que a frequência das palavras na língua poderia ter alguma influência nas trocas. A ideia é que a frequência (na língua) da palavra inserida (falada, lida) deveria ser maior do que a frequência da palavra original (escrita).

Para fazer essa análise, observamos, inicialmente, o panorama geral das trocas de palavra distribuídas por falante, e suas respectivas retratações. Esses dados estão na **Tabela 30**, que descreve também o percentual de trocas por palavra lida e o percentual de retratações por troca de palavra. Note-se que o total de trocas de palavras com prejuízo (284) é menor do que o que aparece em tabelas anteriores (300). Essa diferença resulta da exclusão de trocas de numerais, cuja análise de frequência no Corpus Brasileiro (SARDINHA; MOREIRA FILHO; ALAMBERT, 2019) não poderia ser feita de maneira adequada.

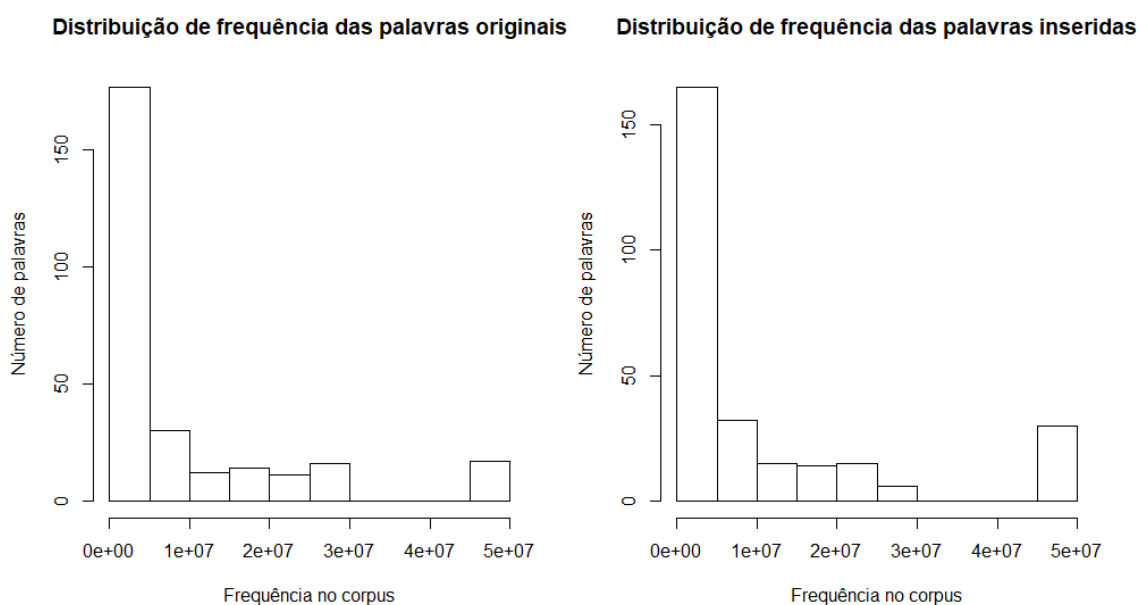
Tabela 30 - Trocas de palavras e retratações por falante

Falante	Palavras lidas	Trocas	Retratações	Troca/palavra lida	Retratação/troca
1	1780	19	5	1.07%	26%
2	259	0	0	0.00%	-
3	795	12	3	1.51%	25%
4	37	1	1	2.70%	100%
5	521	20	7	3.84%	35%
6	901	19	4	2.11%	21%
7	1104	13	0	1.18%	0%
8	1215	8	4	0.66%	50%
9	690	3	0	0.43%	0%
10	1043	10	3	0.96%	30%
11	678	0	0	0.00%	-
12	1607	8	4	0.50%	50%
13	577	2	0	0.35%	0%
14	1224	4	0	0.33%	0%
15	788	14	4	1.78%	29%
16	928	9	3	0.97%	33%
17	913	4	2	0.44%	50%
18	806	9	4	1.12%	44%
19	590	6	4	1.02%	67%
20	1026	4	1	0.39%	25%
21	1386	3	1	0.22%	33%
22	115	7	1	6.09%	14%
23	968	0	0	0.00%	-
24	286	2	1	0.70%	50%
25	1249	2	1	0.16%	50%
26	919	7	5	0.76%	71%
27	2219	14	3	0.63%	21%
28	766	14	2	1.83%	14%
29	1229	11	0	0.90%	0%
30	1072	7	2	0.65%	29%
31	467	5	3	1.07%	60%
32	917	6	2	0.65%	33%
33	638	13	5	2.04%	38%
34	805	10	4	1.24%	40%
35	883	0	0	0.00%	-
36	945	4	3	0.42%	75%
37	355	2	1	0.56%	50%
38	903	9	6	1.00%	67%
39	414	3	0	0.72%	0%
Total	34018	284	89
Mediana	901	0.72%	33%

Dos 39 falantes, apenas quatro não trocaram palavras. As trocas ocorreram em torno de 0,72% do total de palavras lidas. As retratações, por sua vez, corrigiram cerca de um terço das trocas. Em outras palavras, em dois terços dos casos de troca, não houve retratação.

Para avaliar a relação entre troca e frequência da palavra na língua, utilizamos, conforme descrito na metodologia, o número de ocorrências de cada palavra no Corpus Brasileiro (SARDINHA; MOREIRA FILHO; ALAMBERT, 2019). A **Figura 24** mostra os histogramas com o número de ocorrências das palavras originais (presentes no texto) e inseridas (presentes na fala) e suas respectivas ocorrências no Corpus Brasileiro (eixo “x”). As barras verticais, portanto, mostram o número de ocorrências (em nosso corpus) de cada categoria em que dividimos o Corpus Brasileiro (CB), que inclui desde palavras com apenas 1 ocorrência (extremo esquerdo do eixo “x”) até palavras com quase 50 milhões de ocorrências (extremo direito do eixo “x”).

Figura 24 – Histograma de frequências das palavras originais e inseridas



Observa-se que, apesar de os gráficos serem visualmente semelhantes, parece haver uma bscula entre os extremos mais e menos frequentes, com reduo da barra vertical do extremo esquerdo e aumento da barra vertical do extremo direito nas palavras inseridas, em comparao com as originais. Para avaliar estatisticamente essa hiptese alternativa de que as palavras inseridas seriam mais frequentes na lngua (no CB), calculamos a mediana das amostras e aplicamos o teste de Wilcoxon (*signed rank test with continuity correction*). A mediana de ocorrncias (no CB) das palavras originais

foi de 571056, e a das palavras inseridas, de 640246, e o teste de Wilcoxon sugere que as palavras inseridas são realmente mais frequentes no CB, com um *p-value* de 0.03975.

Resolvemos testar também a hipótese de que a retratação possa ter alguma relação com a frequência das palavras originais e inseridas. Ou seja, quando há retratação, a relação de frequências entre palavras originais e inseridas seria diferente de quando não há retratação. O teste de Wilcoxon (*signed rank test with continuity correction*), neste caso, traz resultados interessantes. Quando consideramos as trocas em que não há retratação (dois terços dos casos), a diferença entre as frequências não é significativa (*p-value*=0,58, teste bicaudal). Por outro lado, ao testar apenas o universo das trocas em que há retratação, a diferença é bastante significativa, com medianas de 130640 para as palavras originais e 225027 para as inseridas, e um *p-value* de 0,0057.

Ao observar esses resultados, vemos que se confirma a hipótese de que as palavras inseridas são mais frequentes no CB (e, por extensão, hipoteticamente, na própria língua) do que as originais; mas, curiosamente, essa diferença só se confirma para os casos em que há retratação. Observamos também que as medianas das ocorrências (tanto das palavras originais, quanto das inseridas) são bem menores quando há retratação (originais: 130640; inseridas: 225027), do que quando não há (originais: mediana 1095616, média 9120648; inseridas: mediana 1095616; média 9925219). O teste de Wilcoxon (*rank sum test with continuity correction*) confirma que, no caso das palavras originais, essas diferenças (entre retratadas e não retratadas) são significativas (*p-value* = 0,01301), mas, no caso das inseridas, não (*p-value* = 0,2219). Em resumo, o que observamos foi o seguinte:

- Quando houve troca de palavra **com** retratação, as palavras inseridas eram mais frequentes no CB do que as palavras originais
- Quando houve troca de palavra **sem** retratação, as palavras inseridas tinham a mesma frequência no CB do que as palavras originais
- Além disso, palavras originais que sofreram troca **com** retratação eram **menos** frequentes no CB do que palavras que sofreram troca **sem** retratação
- Ou seja, parece que a frequência das palavras no CB influencia tanto a troca, quanto a retratação: no universo das palavras trocadas, palavras originais menos frequentes são trocadas por palavras mais frequentes, e são retratadas

Esse universo das palavras que sofreram troca não é homogêneo; há palavras que tiveram apenas uma ocorrência de troca, e outras que foram trocadas diversas vezes. A **Tabela 31** mostra, em três colunas adjacentes:

- As 19 palavras originais (presentes no texto escrito) que sofreram mais trocas
- As 19 palavras mais frequentes no conjunto dos 39 textos analisados; ou seja, mais frequentes no nosso corpus (total de 33860 palavras)
- As 19 palavras mais frequentes no CB
- Observação: estão marcadas em cinza as palavras que, mesmo estando entre as mais frequentes nos textos, não sofreram nenhuma troca

Tabela 31 – Número de trocas vs. Ocorrências nos textos vs. Ocorrências no CB

Mais trocadas		Mais frequentes nos textos		Mais frequentes no CB	
Palavra	Trocas	Palavra	Ocorrências	Palavra	Ocorrências
de	17	de	1876	de	45589862
a	16	a	1201	a	28392471
do	14	e	1009	e	23366842
e	7	que	895	o	21888135
da	6	o	902	que	16478129
em	6	do	614	do	15254651
no	6	da	625	da	14011812
com	6	em	541	em	11288244
por	5	para	384	para	8523040
ao	4	com	332	com	7404495
dos	4	no	317	os	7253803
na	4	um	281	no	7072582
para	4	dos	246	um	6039761
o	4	os	279	na	6038903
das	3	uma	224	é	5945831
é	3	é	327	uma	5610681
Nos	3	mais	238	dos	5391117
se	2	na	230	as	5349247
à	2	se	206	não	5123346

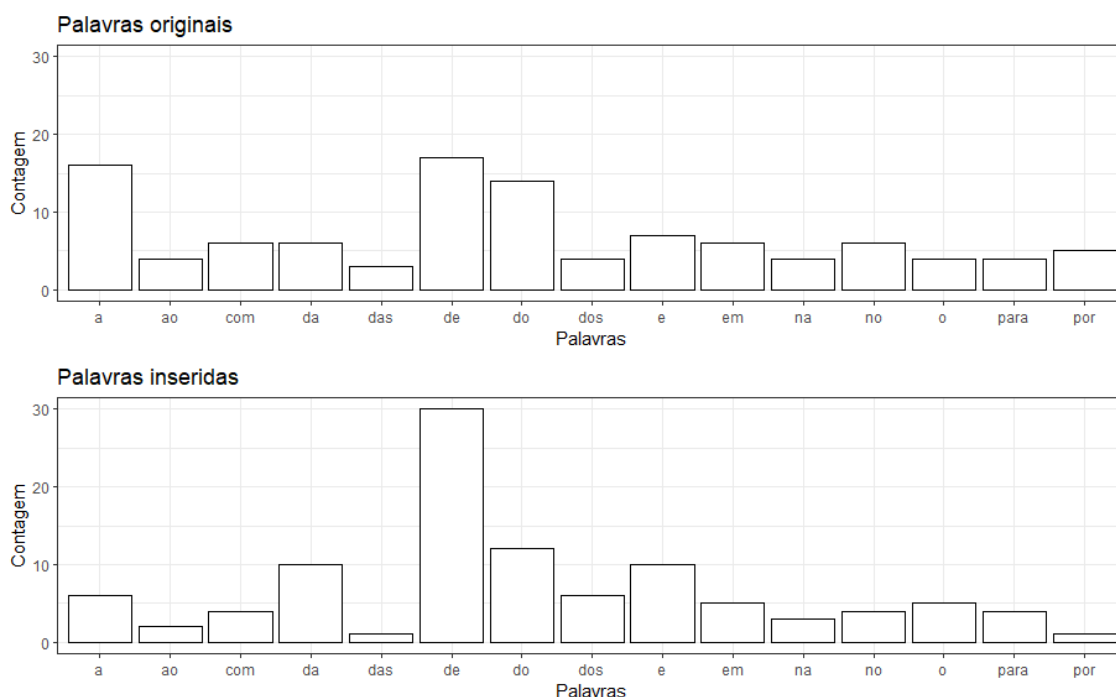
Como se pode ver, há quatro palavras (marcadas em cinza na **Tabela 31**) que, apesar de serem frequentes nos textos escritos (e no CB), não sofreram nenhuma troca. A palavra “que”, por exemplo, é a quarta mais frequente no nosso corpus (textos escritos), e não foi trocada uma vez sequer.

Se aplicarmos o teste do qui-quadrado (*Pearson's Chi-squared test with simulated p-value, based on 2000 replicates*) às ocorrências de trocas vs. ocorrências no corpus (colunas 1 e 2 da **Tabela 31**), obtemos um *p-value* de 0,0004998, mostrando que as distribuições são diferentes. As palavras cujas ocorrências de troca são mais discrepantes de suas ocorrências no corpus foram:

- Palavras que não sofreram troca, apesar estarem entre as mais frequentes no corpus: “que”, “um”, “uma” e “mais”
- Palavras que sofreram trocas desproporcionalmente mais frequentes do que suas ocorrências no corpus: “por”, “ao”, “nos” e “das”

Outra análise que pode ser feita é descobrir se, nos processos de troca de palavra, as ocorrências de determinadas palavras aumentam ou diminuem. Essa análise mostraria se há alguma predileção pela inserção de determinadas palavras, em prejuízo de outras. Por exemplo, a palavra “de” poderia ser mais frequente entre as palavras inseridas do que entre as originais, indicando que ela teria preferência nas “escolhas” feitas pelo leitor; no outro extremo, a palavra “a” poderia ser menos frequente entre as inseridas, indicando que ela estaria sendo preterida nessa “escolha”. A **Figura 25** mostra o número de ocorrências das palavras originais e inseridas mais frequentes.

Figura 25 – Palavras originais e inseridas mais frequentes e o número de vezes em que aparecem nos textos e nas falas lidas



Pode-se observar que, quando ocorrem trocas, algumas palavras têm sua frequência reduzida, e outras, aumentada. O “de”, por exemplo, aumentou bastante sua frequência, enquanto o “a” a reduziu. Entretanto, como o número de ocorrências de cada palavra é relativamente baixo e são poucas as palavras com mudança “significativa” de frequência, os testes estatísticos (*Pearson's Chi-squared test with simulated p-value based on 2000 replicates* e *Fisher's Exact Test for Count Data with simulated p-value based on 2000 replicates, two-sided*) não confirmam a hipótese alternativa de que os processos de troca de palavras implicam redução ou aumento na incidência de determinadas palavras (*p-values* de, respectivamente, 0,3538 e 0,3248).

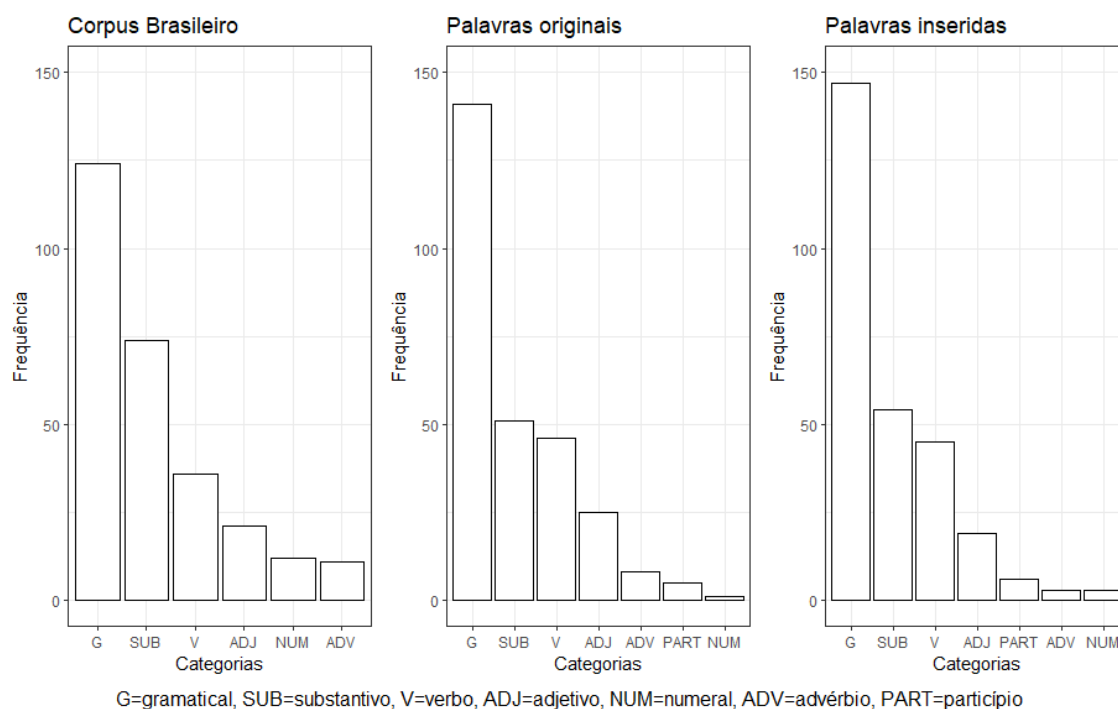
4.2.3.2 As categorias gramaticais das palavras trocadas

Outra abordagem para analisar se as palavras têm características específicas que influenciem a predisposição para troca (substituição da palavra original por uma palavra inserida) seria classificar as palavras trocadas em “palavras lexicais” ou “palavras gramaticais”, conforme detalhado na metodologia.

A **Figura 26** mostra essa distribuição, considerando a ocorrência proporcional de cada categoria entre as palavras originais, as palavras inseridas, e o CB. Com relação ao CB, observe-se

que: a) ele não emprega a classificação “particípio”; b) seu número de ocorrências (na ordem dos milhões) foi proporcionalmente ajustado para a escala do nosso corpus (total de 277 palavras).

Figura 26 – Histogramas mostrando a distribuição das palavras por categoria



Por meio do teste do qui-quadrado (*Fisher's Exact Test for Count Data with simulated p-value based on 2000 replicates*), comparamos a distribuição de frequências das categorias das palavras originais vs. inseridas, considerando como alternativa a hipótese de que as distribuições são diferentes. O resultado não confirma essa hipótese (*p-value* de 0,3323), indicando que as palavras inseridas mantêm a mesma distribuição de categorias lexicais e gramaticais do texto original.

O fato de que as frequências das diversas categorias de palavras são semelhantes entre as palavras originais e inseridas não significa que não haja troca de categoria quando há troca de palavra. A **Tabela 32** mostra o número de palavras de cada categoria entre as palavras originais e inseridas. Observação: excluimos os numerais, que têm apenas uma e três ocorrências, respectivamente.

Tabela 32 - Categorias das palavras originais e inseridas

Categoria	Original	Inserida
adjetivo	25	19
advérbio	8	3
gramatical	141	147
particípio	5	6
substantivo	51	54
verbo	46	45

Observa-se, por exemplo, que, no cômputo geral, alguns adjetivos e advérbios deixaram de ser adjetivos e advérbios no processo de troca de palavra. Os casos individuais de troca de palavra com troca de categoria estão contabilizados na **Tabela 33** e destacados em cinza.

Tabela 33 - Categorias das palavras originais e inseridas

Original	Inserida	Eventos	%
G	G	137	
SUB	SUB	40	
V	V	39	
ADJ	ADJ	9	83%
PART	PART	2	
ADV	ADV	1	
ADJ	SUB	11	
SUB	ADJ	8	
V	G	6	
PART	V	3	
SUB	G	2	
G	V	1	
G	NUM	2	
ADV	SUB	2	
ADV	PART	2	
ADV	ADJ	2	17%
ADJ	PART	2	
V	SUB	1	
SUB	ADV	1	
ADV	G	1	
ADJ	V	1	
ADJ	G	1	
ADJ	ADV	1	
ADV	V	1	

ADJ: adjetivo; ADV: advérbio; SUB: substantivo; NUM: numeral; PART: particípio; V: verbo; G: palavra gramatical

Pode-se ver que houve troca de categoria em 17% dos casos e que as categorias mais trocadas foram adjetivo-> substantivo e substantivo-> adjetivo. Além disso, são bem poucos os casos de troca de categoria entre palavras gramaticais e lexicais e vice-versa. A **Tabela 34** consolida esses dados para evidenciar o percentual de trocas de categoria por cada categoria de palavra trocada.

Tabela 34 - Trocas de palavra com troca de categoria

Categoria	Palavras originais	Troca de categoria na palavra inserida
adjetivo	25	64%
advérbio	8	88%
gramatical	141	3%
particípio	5	60%
substantivo	51	20%
verbo	46	15%

Chamam atenção na **Tabela 34**:

- o percentual de troca de categoria de adjetivos, advérbios e particípios, principalmente os primeiros, que constituem uma amostra com número mais representativo
- e, novamente, a quase ausência de troca de categoria de palavras gramaticais.

A **Tabela 35** mostra alguns exemplos concretos das trocas de palavras em que houve troca de categoria.

Tabela 35 - Palavras originais e inseridas com troca de categoria

Original ADJ	Inserida SUB	Original SUB	Inserida ADJ	Original V	Inserida G
primeiro	presente	política	pública	é	e
próprio	processo	medida	média	é	e
fundamental	fundamento	estratégias	estratégicas	são	dos
burocrática	burocracia	ortopedia	ortopédica	é	e
cortês	cortes	emprestador	espetacular	quero	que
maiores	companhias	princípios	principais		
excelente	excelência				
séria	série				
continentais	continentes				
inspirador	imperador				

4.2.3.3 As semelhanças entre palavra original e inserida

A **Tabela 35** revela outra característica presente na lista de palavras originais e inseridas: muitos pares de palavras original/inserida apresentam semelhança sonora e/ou visual. Ou seja, parece existir uma relação de semelhança fonológica e/ou gráfica entre as palavras trocadas. Para compreender melhor esse fenômeno, analisamos os 277 pares de palavras original/inserida e os classificamos em três categorias, de acordo com o tipo de semelhança, conforme descrito na metodologia:

- **Lexema:** as palavras são variações de forma originadas de um mesmo lexema (p.ex.: de/do, burocracia/burocrático, desenvolve/desenvolveu).
- **Paronímia:** as palavras são semelhantes, mas não são variações de forma originadas de um mesmo lexema (p.ex.: e/é, tem/temo, caos/casos).
- **Nenhuma:** as palavras não apresentam semelhança gráfica/fonológica e não são variações de forma originadas de um mesmo lexema.

A **Tabela 36** mostra a distribuição dessas categorias de semelhança por tipo de palavra (gramatical ou lexical). Palavras lexicais, neste caso, incluem substantivos, adjetivos, verbos, advérbios, numerais e participípios.

Tabela 36 - Semelhança entre palavra original e inserida por tipo de palavra

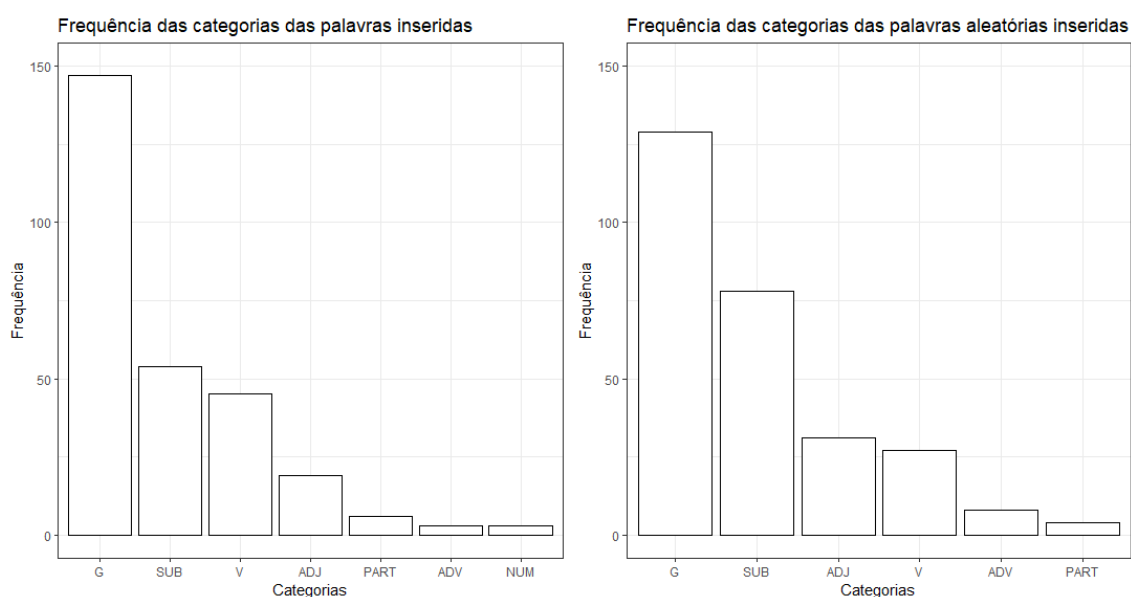
Tipo de palavra	Semelhança			Total	%
	Sim		Não		
	Paronímia	Lexema	Nenhuma		
Gramatical	27(19%)	64(45%)	50(35%)	141	51%
Lexical	75(55%)	50(37%)	11(8%)	136	49%
Total	102(37%)	114(41%)	61(22%)	277	100%
Total	216(78%)		61(22%)		

Observa-se na **Tabela 36** que há uma grande quantidade de palavras com relação de semelhança (78%), seja por paronímia, seja por compartilhamento de lexema. De fato, apenas uma minoria das palavras trocadas (22%) não guarda nenhuma semelhança. Para testar a hipótese de que as trocas de palavra ocorrem, preferencialmente, entre palavras com algum grau de semelhança — ou seja, que a semelhança entre palavra original e palavra inserida não ocorre aleatoriamente —, comparamos os dados da **Tabela 36** com uma lista de palavras aleatórias retiradas do CB. O procedimento utilizado foi o seguinte: para cada palavra original, um algoritmo no Microsoft Excel

escolheu uma palavra aleatória no CB, dentro de uma faixa de frequência próxima à da palavra original. Por exemplo: a palavra “pessoas” teve 497500 ocorrências no CB; o nosso algoritmo selecionou uma palavra aleatória com número de ocorrências entre 447750 e 547250, e selecionou a palavra “vida”. O mesmo procedimento foi repetido para cada palavra original.

As palavras escolhidas aleatoriamente foram classificadas segundo os mesmos critérios utilizados para as palavras originais e inseridas. A **Figura 27** mostra a distribuição dessas palavras (palavras aleatórias) por categorias e a compara com a distribuição das palavras inseridas pelos falantes (palavras inseridas).

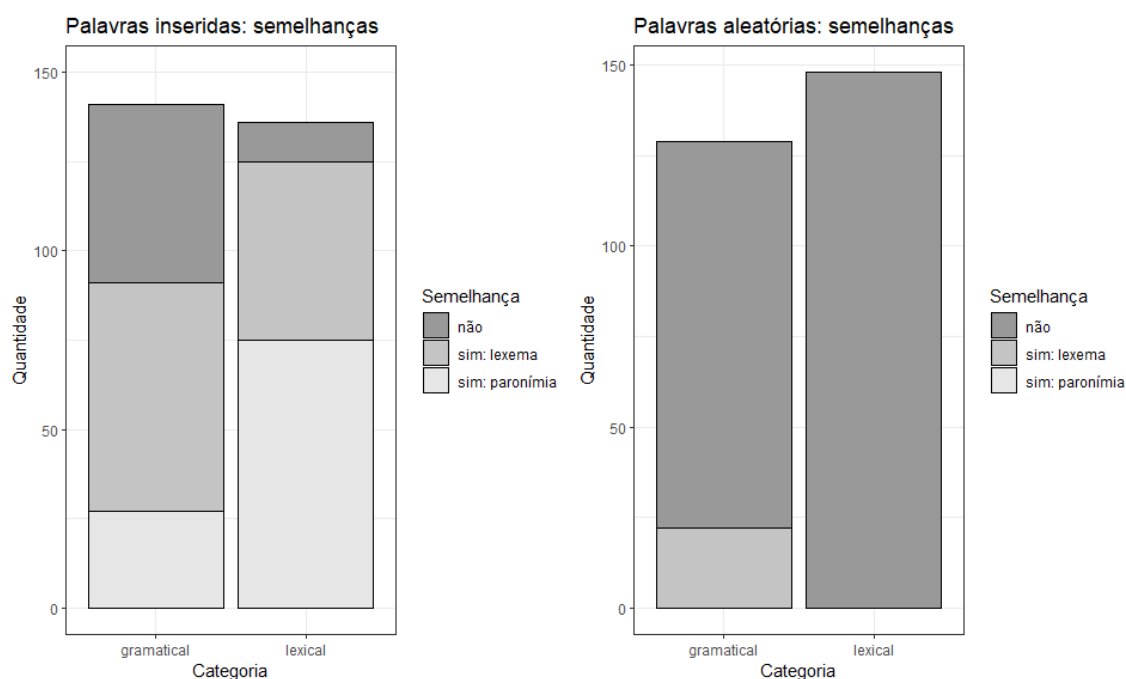
Figura 27 – Palavras aleatórias vs. inseridas: distribuição em categorias



Como se pode ver, a distribuição não é idêntica, mas há consistências, como a maior frequência de palavras gramaticais, os substantivos na segunda posição e os verbos e adjetivos disputando a terceira colocação. Quando analisamos a relação de semelhança entre as palavras originais e aleatórias, observamos que as consistências desaparecem.

A **Figura 28** ilustra a mesma distribuição da **Tabela 36**, substituindo as palavras inseridas (pelos falantes) pelas palavras escolhidas aleatoriamente (pelo algoritmo).

Figura 28 – Semelhanças entre palavras: originais vs. inseridas e originais vs. aleatórias



A **Tabela 37** revela que houve poucos casos de semelhança entre as palavras originais e as palavras escolhidas aleatoriamente no CB. O que vemos são alguns casos de semelhança entre palavras gramaticais que compartilham o mesmo lexema e nenhum caso de semelhança por paronímia, tanto entre palavras gramaticais, quanto lexicais.

Tabela 37 - Relação entre palavra original e aleatória* por tipo de palavra

Tipo de palavra	Relação entre palavras			Total
	Paronímia	Lexema	Nenhuma	
Gramatical	0	22	107	129
Lexical	0	0	148	148
Total	0	22	255	277
%	0%	8%	92%	-

*escolhida aleatoriamente no CB dentro uma faixa de frequência próxima à da palavra original

A **Figura 28** ilustra graficamente as diferenças de proporção das semelhanças entre palavras originais e palavras inseridas, e entre palavras originais e palavras aleatórias. Para confirmar essa observação de que as semelhanças entre palavras originais e inseridas não são fruto do acaso, aplicamos o teste do chi-quadrado (Pearson's Chi-squared test with Yates' continuity correction) aos dados da **Tabela 36** e da **Tabela 37**, comparando as semelhanças entre os pares de palavras original/inserida vs. original/aleatória, e obtivemos os seguintes resultados:

- a) Palavras gramaticais: $p\text{-value} = 7,429e-15$.
- b) Palavras lexicais: $p\text{-value} < 2,2e-16$.
- c) Todas as palavras (gramaticais e lexicais): $p\text{-value} < 2.2e-16$.

Esses resultados mostram que, de fato, as palavras inseridas são semelhantes, fonologicamente (por paronímia ou compartilhamento de lexema), às palavras originais; e que esta semelhança não ocorre aleatoriamente, pois uma seleção aleatória de palavras de frequência semelhante retiradas do CB não retorna, via de regra, palavras fonologicamente semelhantes às originais.

4.2.4 Resumo da seção 4.2 (incongruências segmentares)

- Incongruências mais comuns (três quartos dos casos) e respectivas retratações:
 - Desvios de ortoépia (metade é retratado)
 - Trocas de palavras (um terço é retratado)
 - Desvios de concordância (um décimo é retratado)
- Sobre o universo das palavras trocadas:
 - Retratação associa-se a:
 - Palavra original menos frequente no CB
 - Palavra inserida mais frequente no CB do que a original
 - Palavras com incidência de trocas desproporcional à sua frequência:
 - Nunca trocadas: que, um, uma, mais
 - Mais trocadas do que “deveriam”: por, ao, nos, das
 - Sobre categorias lexicais:
 - 17% das trocas envolvem troca de categoria lexical
 - Mas as palavras inseridas mantêm a mesma distribuição de categorias das palavras originais
 - Quase não há troca de palavra lexical por gramatical e vice-versa
 - Troca de categoria mais comum: adjetivo -> substantivo
 - Sobre semelhança entre palavra original e inserida
 - Quase 80% das palavras inseridas são semelhantes às palavras originais

4.3 As incongruências suprasegmentares

A descrição e análise das 858 ocorrências de incongruências suprasegmentares organiza-se da seguinte forma:

- 4.3.1 listamos as classificações das incongruências suprasegmentares e fornecemos exemplos concretos retirados do corpus
- 4.3.2 explicamos como foi realizado o teste Kappa de Cohen para validação do julgamento da inserção/deleção incongruente de fronteiras e descrevemos e analisamos seus resultados
- 4.3.3 mostramos o número de ocorrências das diversas incongruências suprasegmentares identificadas no corpus
- 4.3.4 analisamos a relação entre as incongruências suprasegmentares e as seguintes “variáveis independentes”:
 - 4.3.4.1 incongruências vs. segmentos do discurso, conforme Grosz e Sidner (1986);
 - 4.3.4.2 incongruências vs. “estratégia de identificação do ator” (*actor identification strategy* – AIS), descrita na metodologia do trabalho
- 4.3.5 resumimos os achados da seção 4.3

4.3.1 Descrição das incongruências suprasegmentares

A análise do corpus revelou um total de 858 eventos de incongruências suprasegmentares. O sistema de classificação dessas incongruências e alguns exemplos concretos retirados do corpus são fornecidos abaixo (a localização das incongruências está identificada com uma barra inclinada – “/”):

- Pausa: “Quem não a acompanha, quem / não contribuiu e não participa da marcha, acaba / esquecido no acostamento de uma estrada qualquer.”
- Inserção de fronteira não terminal (INT): “A questão é voltar a gerar empregos; é dar trabalho / e devolver a esperança aos brasileiros...”

- Inserção de fronteira terminal (IT): “...é sintoma de um problema muito mais profundo, de uma doença muito mais séria e mais grave que põe em xeque a própria legitimidade da classe política / e da democracia brasileira.”
- Deleção de fronteira não terminal (DNT): “Esse respeito e essa confiança — não nos enganemos sobre isso — jamais serão recuperados por meio da doutrinação, / do marketing, ou dos currículos escolares.”
- Hesitação (Hes): “Sou o mesmo c/idade que...”
- Deleção de fronteira terminal (DT): “É certo que, em 2015, não havia crise, já que os reservatórios estavam com 100% de suas capacidades. / Durante o ano de 2016, contudo, o brasiliense viu menos água que o comum...”
- Lentificação (Lent): “Cui/dado com o meio ambiente, uso moderado da água, repressão dos casos de ilegalidades são medidas necessárias e urgentes.”
- Afirmação lida como pergunta: autoexplicativa
- Pergunta lida como afirmação: autoexplicativa
- Outros: eventos de incongruência prosódica não enquadrados nas classificações anteriores

A classificação “outros” incluiu manifestações prosódicas que causaram estranheza perceptiva ao pesquisador. Via de regra, apresentou entoação que pareceu inadequada ao texto lido, mas que não se enquadrou nas demais classificações. Em virtude dessa indefinição, optamos por agrupá-las em uma categoria menos conceitual. Por exemplo:

- “Desde a sua fundação, a TV Record vivenciou diversos momentos marcantes /, momentos de glórias e de lutas essenciais para o seu futuro.”
 - Descrição: entoação (perceptivamente) ascendente, sugerindo para o pesquisador que o leitor iria dizer, por exemplo: “momentos marcantes e de lutas essenciais para seu futuro.”
- “No limite, / essas taxas de juros podem até levar o sistema financeiro ao colapso.”
 - Entoação (perceptivamente) descendente, quando se esperaria uma entoação ascendente.

4.3.2 Validação das incongruências do tipo inserção/deleção de fronteiras

Um conjunto de classificações que mereceu uma atenção especial foi o que envolveu inserções e deleções de fronteiras prosódicas. Trata-se de um grupo mais bem definido do que o grupo “outros”; mas não tão bem definido e de identificação tão simples quando as “pausas”, “hesitações” e “lentificações”. Esse grupo (inserções/deleções de fronteiras) mereceu um estudo paralelo, com a intenção de validar o julgamento do pesquisador, conforme descrito na metodologia. Recrutamos quatro voluntários, falantes de PB, naturais de Brasília (DF), com as seguintes características demográficas:

Voluntário 1. 18 anos, sexo masculino, estudante de Ciência da Computação;

Voluntário 2. 22 anos, sexo feminino, estudante de Biotecnologia;

Voluntário 3. 20 anos, sexo masculino, estudante de Engenharia;

Voluntário 4. 16 anos, sexo masculino, estudante do Ensino Médio.

Esses voluntários participaram de um experimento que consistiu nos seguintes passos:

1. Foram informados de que:

- a. as únicas informações relativas ao experimento que seriam consideradas e divulgadas seriam dados numéricos ou genéricos, sem identificação pessoal
- b. tratava-se, linhas gerais, de uma pesquisa envolvendo incongruências prosódicas de leitura
- c. com relação às incongruências, não haveria resposta certa ou errada, pois se tratava, apenas, de validar, ou não, o julgamento do pesquisador

2. Foram apresentados a quatro períodos com incongruências do tipo inserção ou deleção de fronteira, retirados aleatoriamente do corpus, sendo que a apresentação de cada período respeitou a seguinte ordem:

- a. Leitura do período com a escrita original

- b. Audição do áudio de leitura do período
 - c. Leitura do período com marcações indicando inserções (/) ou deleções (*) incongruentes de fronteira, segundo o julgamento do pesquisador
 - d. Repetição de qualquer dos passos anteriores, a critério do voluntário
3. Cada voluntário recebeu um arquivo contendo o texto escrito e o áudio de leitura de 40 pares de períodos retirados do corpus, cada par consistindo em:
- a. um período com incongruência do tipo inserção ou deleção de fronteira
 - b. um período congruente
4. Cada voluntário foi orientado a escolher, no mínimo, 10 pares de períodos e:
- a. ignorar qualquer incongruência que não pudesse ser considerada inserção ou deleção de fronteira;
 - b. selecionar, em cada par, um período em que ele julgasse haver alguma incongruência do tipo inserção ou deleção de fronteira;
 - c. identificar, no período considerado incongruente, o local da inserção ou deleção inadequada de fronteira prosódica.

Os julgamentos dos voluntários quando à identificação do período incongruente em cada par de períodos (período 1 ou período 2) foram confrontados com os julgamentos do pesquisador. Aplicou-se o teste Kappa de Cohen, considerando como hipótese alternativa a existência de concordância entre os julgamentos dos voluntários e do pesquisador, com um intervalo de confiança de 95%. A **Tabela 38**, adaptada de Landis e Koch (1977), descreve os possíveis graus de concordância a serem obtidos pelo teste:

Tabela 38 - Grau de concordância entre pesquisador e voluntários

Kappa	Grau de concordância
<0,00	Pobre (<i>Poor</i>)
0,00 - 0,20	Leve (<i>Slight</i>)
0,21 - 0,40	Razoável (<i>Fair</i>)
0,41 - 0,60	Moderado (<i>Moderate</i>)
0,61 - 0,80	Substancial (<i>Substantial</i>)
0,81 - 1,00	Quase perfeito (<i>Almost perfect</i>)

A **Tabela 39** mostra os resultados do teste Kappa de Cohen para cada um dos voluntários.

Tabela 39 - Resultados dos graus de concordância entre pesquisador e voluntários

Teste Kappa de Cohen				
Voluntário	Kappa	Grau de concordância	Intervalo de confiança	Pares de períodos avaliados
1	0,69	Substancial	0,42-0,95	25
2	0,61	Substancial	0,12-1,00	11
3	0,44	Moderado	0,03-0,86	18
4	0,27	Razoável	(-)0,19-0,73	13

Pode-se observar que, apesar de se tratar de voluntários sem conhecimento linguístico acadêmico, instruídos de forma sumária e superficial sobre o fenômeno das fronteiras prosódicas, houve um grau de concordância entre os julgamentos do pesquisador e dos voluntários que variou de razoável a substancial. Tendo em vista que o voluntário 1 empenhou-se mais em sua tarefa, analisando um número maior de pares de sentenças e obtendo um grau de concordância maior do que os demais, resolvemos aplicar o teste Kappa de Cohen às marcações dos locais onde o voluntário julgou haver inserção ou deleção inadequada de fronteira prosódica, confrontando suas marcações com as marcações do pesquisador. A **Tabela 40** mostra o resultado do teste.

Tabela 40 - Concordância entre o pesquisador e o voluntário 1 na marcação do local de fronteiras inadequadas

Voluntário	Kappa	Grau de concordância	Intervalo de confiança	Marcações avaliadas
1	0,43	Moderado	0,27-0,58	40

Percebe-se que, além de concordar substancialmente com o pesquisador quanto à percepção de incongruência de leitura, o voluntário também teve um julgamento semelhante com relação à

localização e ao tipo de inadequação prosódica de fronteira. A partir desse resultado, consideramos estatisticamente validado o julgamento de fronteiras prosódicas do pesquisador e prosseguimos com a descrição e análise das incongruências suprasegmentares.

4.3.3 Descrição geral das incongruências suprasegmentares identificadas

A **Tabela 41** descreve o universo dos eventos de incongruências suprasegmentares anotados, classificados e distribuídos por falante. As marcações em cinza indicam que o falante não produziu determinado tipo de incongruência. Empregam-se, também, as seguintes convenções: INT: inserção de fronteira não terminal; IT: inserção de fronteira terminal; DNT: deleção de fronteira não terminal; Hes: hesitação; DT: deleção de fronteira terminal; Lent: lentificação; A-> p: leu afirmação como pergunta; P-> a: leu pergunta como afirmação.

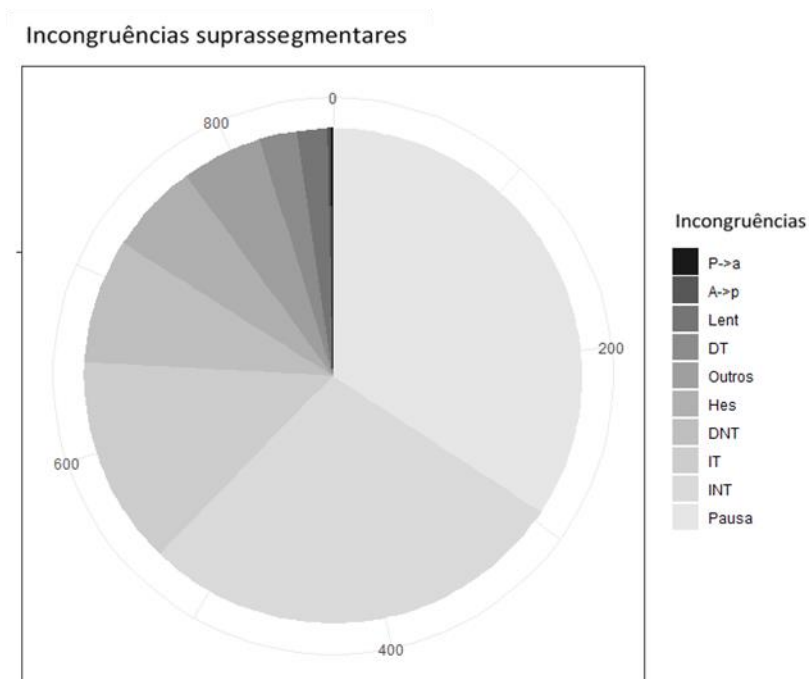
Tabela 41 - Incongruências suprasegmentares por falante

Ordenação vertical: total por falante - Ordenação horizontal: total por tipo de incongruência

Falante	Pausa	INT	IT	DNT	Hes	Outros	DT	Lent	A-> p	P-> a	Total
36	26	10	0	7	2	3	1	0	0	0	49
27	6	23	5	8	3	1	1	0	0	0	47
18	20	11	3	4	5	3	0	0	0	0	46
29	13	16	7	1	0	4	2	1	0	0	44
35	19	12	3	0	3	2	0	2	1	0	42
8	13	9	9	6	1	1	0	1	0	0	40
3	6	19	5	7	0	1	0	1	0	1	40
6	11	10	3	2	2	3	5	0	0	0	36
33	13	12	3	1	1	1	0	3	0	0	34
21	8	13	4	3	0	5	0	0	0	0	33
26	10	9	3	2	1	2	1	3	0	0	31
10	12	6	8	2	1	0	0	0	0	0	29
5	10	4	5	4	2	2	0	1	0	0	28
19	7	11	3	2	0	2	0	0	1	0	26
7	7	9	1	4	2	2	0	1	0	0	26
38	11	2	6	1	4	1	0	0	0	0	25
17	7	5	6	5	0	1	1	0	0	0	25
15	14	4	4	2	1	0	0	0	0	0	25
28	11	4	1	0	5	1	1	1	0	0	24
16	4	8	5	1	0	0	4	1	0	0	23
30	12	2	2	1	2	1	0	0	0	0	20
39	4	7	4	0	1	1	0	0	0	0	17
20	11	2	1	0	2	0	0	0	0	0	16
14	2	5	3	2	0	0	3	0	0	0	15
32	3	5	1	0	2	1	0	2	0	0	14
1	8	1	1	1	3	0	0	0	0	0	14
25	0	6	1	1	2	2	0	0	0	0	12
13	4	5	0	1	1	1	0	0	0	0	12
12	3	1	6	0	1	1	0	0	0	0	12
34	1	0	5	0	3	0	0	0	0	0	9
31	4	1	0	1	1	0	1	0	0	0	8
11	2	2	3	0	0	0	1	0	0	0	8
24	2	2	2	0	0	0	0	0	0	0	6
9	3	1	0	0	0	2	0	0	0	0	6
2	3	2	1	0	0	0	0	0	0	0	6
37	1	1	2	0	0	0	0	0	0	0	4
22	3	0	1	0	0	0	0	0	0	0	4
23	0	0	0	1	0	1	0	0	0	0	2
Total	294	240	117	70	51	45	21	17	2	1	858
%	34,3%	28,0%	13,6%	8,2%	5,9%	5,2%	2,4%	2,0%	0,2%	0,1%	100%

Uma representação visual da última linha da **Tabela 41** (totais por tipo de incongruência) está ilustrada no gráfico de setores da **Figura 29**.

Figura 29 – Incongruências suprasegmentares (ordem crescente de frequência)



O que observamos é que há incongruências suprasegmentares bastante comuns, como as pausas e as inserções de fronteira, e outras bem raras, como troca de tipologia frasal (afirmativa/interrogativa), lentificação e deleção de fronteira terminal. Ao contrário do que ocorreu com as incongruências segmentares, houve apenas 4 retratações de incongruências suprasegmentares, 2 em INT, e 2 em DNT.

4.3.4 Incongruências suprasegmentares vs. “variáveis independentes”

4.3.4.1 Incongruências suprasegmentares vs. segmentos do discurso

Uma das análises que propusemos dizia respeito aos segmentos de discurso, segundo Grosz e Sidner (1986). Identificamos, para cada período, sua relação de continuidade temática com o período anterior.

Na prática, cada período (considerado como um segmento do discurso, segundo nosso protocolo), classificou-se e rotulou-se de acordo com os seguintes critérios:

- DS 1:
 - tópico do período atual (ou da oração principal do período atual) não tem relação com o tópico ou foco do período anterior
- DS 2:
 - tópico do período atual (ou da oração principal do período atual) foi tópico ou foco no período anterior
 - tópico do período atual (ou da oração principal do período atual) é a primeira pessoa (considerado como tópico ou foco em qualquer momento do discurso)

Para ampliar um pouco o escopo da análise, incluímos também, num primeiro momento, as incongruências segmentares, avaliando a relação entre DS e todo o universo de incongruências analisadas. Os resultados são descritos na **Tabela 42**. Lembrando que “I” indica “incongruente” e “C” indica “congruente”; e que a célula no cruzamento da linha “DS 2” com a coluna “Períodos I”, por exemplo, anota o número de ocorrências de períodos incongruentes que foram classificados como DS 2 (o tópico deste período é o mesmo tópico ou foco do período anterior).

Tabela 42 - Incongruências vs. *discourse segments*

	Períodos I	Períodos C	Total
DS 2	476	320	796
DS 1	375	252	627

DS = *discourse segment*; 2 = continuidade de tópico; 1 = descontinuidade de tópico

O teste do qui-quadrado (*Pearson's Chi-squared test with Yates' continuity correction*) mostra que a proporção de períodos DS 2 e DS 1 é a mesma entre os períodos congruentes e incongruentes; ou seja, na amostra analisada, não houve relação entre o tipo de DS e a ocorrência de incongruências ($p\text{-value}=1$).

Na sequência, analisamos exclusivamente a relação entre continuidade temática (DS1/DS2) e a incongruência de fronteiras prosódicas. Períodos em que houve inserções/deleções incongruentes de fronteiras prosódicas foram classificados como incongruentes (I) e os demais, como congruentes (C). Na **Tabela 43**, podemos observar um número desproporcionalmente alto de períodos tematicamente contínuos (DS2) entre os períodos incongruentes. O teste do qui-quadrado (*Pearson's Chi-squared test with Yates' continuity correction*) confirma que essa desproporção é significativa ($p\text{-value} =$

0,0006345), sugerindo associação entre continuidade temática e incongruência de fronteiras prosódicas.

Tabela 43 – Incongruências de fronteiras prosódicas vs. continuidade temática

Continuidade temática	Períodos (n(%))			Incongruências (n)
	D	F	Total	
DS2	223(73%)	320(56%)	543	264
DS1	81(27%)	252(44%)	333	184
Total(%)	304(100%)	572(100%)	876	448

Para compreender melhor a correlação entre DS e incongruências de fronteiras prosódicas, segregamos os números da **Tabela 43** para obter dados relativos aos subtipos de inserção/deleção de fronteiras. A **Tabela 44** mostra esses dados.

Tabela 44 - Subtipos de inserção/deleção de fronteiras vs. *discourse segments*

	INT	IT	DNT	DT
DS 2	151	62	38	13
DS 1	89	55	32	8

INT: inserção de fronteira não terminal; IT: inserção de fronteira terminal; DNT: deleção de fronteira não terminal; DT: deleção de fronteira terminal

Notamos que a proporção DS1/DS2 nas inserções de fronteira não terminal (INT) foi de 89/151, sugerindo uma associação entre INT e DS2, mas tal associação não atinge um nível de significância estatística (Pearson's Chi-squared p -value = 0,1426).

4.3.4.2 Incongruências suprasegmentares vs. “estratégia de identificação do autor”

Analisamos também os atores das orações contidas nos períodos dos textos escritos e correlacionamos seu grau de “agentividade” com a presença ou ausência de incongruências suprasegmentares do tipo inserção/deleção de fronteiras. A classificação dos atores (ou sujeitos, especialmente quando se tratar de voz passiva) inspirou-se na estratégia de identificação de ator (“*actor identification strategy* – AIS”), descrita na metodologia. As classificações (e convenções) utilizadas foram as seguintes:

- Sujeito

- Voz ativa: A
- Voz passiva: V
- Desinencial: D
- Coordenado (mais de um sujeito no mesmo verbo): C
- Posição:
 - Antes do verbo: 1
 - Depois do verbo: 0
- Pessoa:
 - Primeira pessoa: 1
 - Demais pessoas: 0
- Humano:
 - Sim: 1
 - Não: 0
- Animado:
 - Sim: 1
 - Não: 0
- Definido:
 - Sim: 1
 - Não: 0

Abaixo, são fornecidos alguns exemplos concretos (retirados do corpus) de classificações dos atores de algumas orações:

- “A **notícia** ilustra o nível técnico a que chegou a **IBM**”
 - notícia: A10001 – voz ativa, posição 1, pessoa 0, humano 0, animado 0, definido 1
 - IBM: A00001 – voz ativa, posição 0, pessoa 0, humano 0, animado 0, definido 1
- “Tais **alterações** viabilizaram-se mediante a aprovação do Projeto de Lei...”
 - alterações: V10001 – voz passiva, posição 1, pessoa 0, humano 0, animado 0, definido 1
- “Por esta razão, **optei** por perfilar-me ao lado dos agentes...”
 - optei (verbo indicando o sujeito desinencial): D11111 – sujeito desinencial (eu), posição 1, pessoa, 1, humano 1, animado 1, definido 1

- “Estiveram lá as **senhoras** e os **senhores...**”
 - senhoras: A00111 – voz ativa, posição 0, pessoa, 0, humano 1, animado 1, definido 1
 - senhores: AC00111 – voz ativa, sujeito coordenado, posição 0, pessoa, 0, humano 1, animado 1, definido 1

A **Tabela 45** demonstra as classificações mais comuns (em um universo de 38 classificações diferentes identificadas no corpus) em períodos congruentes (C) e incongruentes (I), considerando, como já dissemos, apenas as incongruências causadas por inserção/deleção de fronteiras.

Tabela 45 - Classificação dos atores (AIS) em períodos congruentes (C) e com incongruências (I) do tipo inserção/deleção de fronteira

Classificação	Períodos C	%	Períodos I	%
A10001	326	46%	198	40%
D11110	93	13%	48	10%
D11111	65	9%	24	5%
D10001	30	4%	33	7%
A10000	27	4%	31	6%
A00001	23	3%	16	3%
A11111	22	3%	5	1%
A10110	21	3%	17	3%
A10111	20	3%	31	6%
V10001	19	3%	17	3%
A00000	10	1%	7	1%

Obs: as 10 classificações mais comuns de 38 que ocorreram

Observa-se que o tipo de ator mais comum é um sujeito de voz ativa, situado antes do verbo, não primeira pessoa, não humano, não animado, e definido, como ilustrado no exemplo fornecido anteriormente: “A **notícia** ilustra o nível técnico...”. O segundo, o terceiro e o quarto tipos de atores mais comuns (totalizando, junto com o primeiro colocado, mais de 60% dos sujeitos) são sujeitos de voz ativa, desinenciais, anteriores ao verbo, como no exemplo: “...**optei** por perfilar-me ao lado dos agentes...”.

O teste do qui-quadrado (*Pearson's Chi-squared test with simulated p-value*) aplicado à tabela completa de classificação de atores (38 entradas) sugere que a distribuição de classificações dos atores é diferente entre as sentenças I e C ($p\text{-value} = 0,0004998$), o que indica haver alguma correlação entre o tipo de ator (conforme a classificação que adotamos) e a incidência ou não de incongruência do tipo

inserção/deleção de fronteiras. Aplicando o teste apenas à **Tabela 45** (as 10 classificações mais comuns), o resultado é ainda mais significativo ($p\text{-value} = 0,0002332$).

Para identificar quais tipos de sujeito poderiam ser responsáveis pela significância identificada pelo teste do qui-quadrado, ordenamos os dados da **Tabela 45** de acordo com as razões entre o percentual de cada classificação em cada tipo de período (C ou I). A **Tabela 46** mostra esse ordenamento, sendo que a coluna “Razão” exibe o resultado da divisão entre os percentuais das colunas “Períodos C” e “Períodos I” sempre dividindo o **maior** pelo **menor** valor. O objetivo dessa metodologia é evidenciar as classificações em que há maior divergência entre períodos congruentes e incongruentes, independentemente da direção (congruência ou incongruência).

Tabela 46 - Dados da **Tabela 45** ordenados pela razão entre Períodos I e Períodos C

Classificação	Períodos C	Períodos I	Razão
A11111	3%	1%	3.09
A10111	3%	6%	2.21
D11111	9%	5%	1.90
A10000	4%	6%	1.64
D10001	4%	7%	1.57
D11110	13%	10%	1.36
V10001	3%	3%	1.28
A10001	46%	40%	1.15
A10110	3%	3%	1.15
A00001	3%	3%	1.01
A00000	1%	1%	1.00

De acordo com a **Tabela 46**, percebemos que:

- A11111 é 3 vezes mais frequente em C
- A10111 é 2 vezes mais frequente em I
- D11111 é quase 2 vezes mais frequente em C, e assim por diante

Para compreender melhor esse fenômeno, segregamos e agrupamos os elementos da classificação de duas formas diferentes e os submetemos testagens estatísticas distintas. O primeiro agrupamento considerou apenas a classificação em voz ativa/passiva e sujeito desinencial, desconsiderando as demais classificações, e está ilustrado na **Tabela 47**.

Tabela 47 - Sujeitos desinenciais ou de voz ativa ou passiva em períodos congruentes (C) e com incongruências (I) do tipo inserção/deleção de fronteira

Classificação	Períodos C	%	Períodos I	%
Ativa	468	66%	334	67%
Passiva	44	6%	37	7%
Desinencial	198	28%	127	26%

O teste do chi-quadrado (*Pearson's Chi-squared test*) não sugere diferença entre as distribuições desses tipos de sujeito entre períodos I e C ($p\text{-value} = 0,5115$).

A segunda forma de segregar e reagrupar as classificações foi desconsiderar as categorias da **Tabela 47** e considerar apenas as classificações dos tipos posição, pessoa, humano, animado e definido. A **Tabela 48** mostra a distribuição da incidência dessas classificações em períodos I e C ordenando-as pelas razões entre os percentuais de ocorrência. Notar que essas razões indicam a magnitude da diferença entre os percentuais de ocorrência, mas não sua direção (para maior congruência, ou incongruência)

Tabela 48 - Classificação dos sujeitos em períodos C e I

Classificação	Períodos C	%	Períodos I	%	Razão
10111	27	4%	43	9%	2.27
11111	88	12%	30	6%	2.06
10000	35	5%	45	9%	1.83
10110	24	3%	25	5%	1.49
11110	93	13%	50	10%	1.30
00000	18	3%	11	2%	1.15
00001	35	5%	28	6%	1.14
10001	384	54%	258	52%	1.04
00111	3	0%	2	0%	-
01110	0	0%	1	0%	-
01111	1	0%	2	0%	-
00110	2	0%	3	1%	-

Obs: desconsiderando a classificação em voz ativa/passiva ou sujeito desinencial

Ignorando as classificações com baixa ocorrência, pode-se observar um contraste entre as duas primeiras colocadas da **Tabela 48**. A classificação 10111 (posição antes do verbo, não primeira pessoa, humano, animado e definido) associa-se desproporcionalmente a períodos I, enquanto a classificação 11111, que se diferencia da anterior apenas no item “pessoa” (trata-se de primeira pessoa), associa-se preferencialmente a períodos C.

O teste qui-quadrado (*Pearson's Chi-squared test with simulated p-value*) aplicado aos dados da **Tabela 48** indica que as diferenças observadas são significativas ($p\text{-value} = 0,0004998$), sugerindo que, de alguma forma, algum aspecto dessas diversas classificações correlaciona-se com a presença ou ausência de incongruências do tipo inserção/deleção de fronteira.

Para tentar identificar qual ou quais das diversas classificações descritas na **Tabela 48** são responsáveis pelas diferenças observadas, segregamos/agrupamos os dados de acordo com cada unidade classificatória (posição, pessoa, humano, animado, definido). A **Tabela 49** ilustra os dados resultantes dessa segregação/agrupamento.

Tabela 49 - Segregação/agrupamento dos dados da **Tabela 28**

Posição	F	D
0	59	48
1	651	451
Pessoa	F	D
0	528	416
1	182	83
Humano	F	D
0	472	342
1	238	157
Animado	F	D
0	472	342
1	238	157
Definido	F	D
0	172	136
1	538	363

A primeira coisa que notamos na **Tabela 49** é que os dados relativos a “humano” e “animado” são idênticos. De fato, revendo os dados, percebemos que não há, nos textos escritos, nenhuma referência a animais ou outras entidades não humanas animadas que estejam na posição de sujeito. Sendo assim, todos os sujeitos classificados como “animados” são também “humanos”.

A segunda observação diz respeito à ocorrência de sujeitos de primeira pessoa em períodos congruentes. Para testar essa discrepância e, adicionalmente, todos os demais dados da **Tabela 49**, aplicamos o teste do qui-quadrado (*Pearson's Chi-squared test with Yates' continuity correction*) a cada uma das subtabelas.

Os resultados de $p\text{-value}$ para cada subclassificação foram os seguintes: definido (0,2614); humano (0,4909); posição (0,4925); pessoa (0,0002584). Tais resultados sugerem que há algum tipo

de associação entre sujeito em primeira pessoa e períodos congruentes. Observando novamente a **Tabela 48**, pode-se extrapolar essa associação e levantar a hipótese de que sujeitos em primeira pessoa em posição pré-verbal associam-se a períodos congruentes. Para testar essa hipótese, agrupamos os dados da **Tabela 49** de acordo com as subclassificações “pessoa” e “posição” e organizamos esses dados na **Tabela 50**.

Tabela 50 - Pessoa/Posição vs. Congruência

Classificação	Períodos C	Períodos I
00	58	45
01	1	3
10	470	371
11	181	80

00: posição pós-verbal + não primeira pessoa; 01: posição pós-verbal + primeira pessoa; 10: posição pré-verbal + não primeira pessoa; 11: posição pré-verbal + primeira pessoa

Aplicando o teste de Fisher (*Fisher's Exact Test for Count Data*) aos dados da **Tabela 50** — considerando, separadamente, as duas primeiras linhas (posição pós-verbal) e as duas últimas linhas (posição pré-verbal) —, obtemos, respectivamente, os *p-values* de 0,3236 e 0,00009917. Devemos notar que há pouquíssimas ocorrências de sujeito em primeira pessoa na posição pós-verbal (segunda linha da **Tabela 50**), o que não nos permite comparar a influência da posição sobre a associação entre primeira pessoa e congruência. Dessa forma, o que podemos dizer é que existe uma associação entre sujeito pré-verbal em primeira pessoa (p.ex.: “Eu sou congruente”) e congruência prosódica de fronteiras, ou seja, ausência de inserção/deleção imprópria de fronteiras prosódicas.

4.3.5 Resumo da seção 4.3

- Descrição geral
 - Pausas e inserções de fronteira prosódica compõem 75% das incongruências suprasegmentares
 - Retratações são raríssimas
- Análises
 - Segmentos do discurso
 - A continuidade temática associou-se a incongruência de fronteiras prosódicas, de um modo geral.
 - “Estratégia de identificação do ator” (AIS)
 - Sujeitos mais comuns (60% dos casos):

- voz ativa, anterior ao verbo, não primeira pessoa, não humano, não animado, definido. P.ex.: “A notícia ilustra o nível técnico...”.
- voz ativa, anterior ao verbo, desinencial. P.ex: “...optei por perfilar-me ao lado dos agentes...”
- Associação entre tipo de ator e congruência/incongruência
 - Primeira pessoa associa-se a congruência

4.4 Predicação primária: palavras/constituintes obrigatórios vs. opcionais

Os períodos congruentes e incongruentes (por inserção/deleção de fronteiras prosódicas) tiveram suas palavras gráficas classificadas segundo critérios de hierarquia relativos ao que chamamos de “predicação primária”, conforme descrito na Metodologia, subseção 3.8.6. Cada palavra foi classificada com um número que variou de 1 a 16, conforme sua classificação hierárquica dentro do período.

O número 1 corresponde ao primeiro nível na hierarquia, indicando que a palavra compõe a predicação primária, indispensável à compreensão semântica mínima do período. Isso significa que essa palavra (de nível 1): a) pertence a um constituinte obrigatório; b) tem presença obrigatória dentro desse constituinte; c) é parte indispensável da predicação primária. Retomando os conceitos utilizados na Metodologia, a palavra de hierarquia 1 é indispensável à **autonomia pragmática** e à **interpretabilidade em isolamento** do segmento de período de que faz parte; e esse segmento de período, por sua vez, corresponde a um **enunciado em potencial**; e, além disso, é o primeiro enunciado em potencial a surgir no período analisado.

As demais classificações (2 a 16) indicam palavras que poderiam ter sido omitidas na leitura sem que o período perdesse sua compreensibilidade mínima; ou seja, elas não pertencem à predicação primária. Note-se que o valor máximo encontrado, de número “16”, não foi determinado arbitrariamente, mas correspondeu, simplesmente, à hierarquia “mais baixa” observada no corpus analisado.

A **Tabela 51** mostra o número de ocorrências (palavras) em cada nível de classificação hierárquica nos períodos congruentes e incongruentes (onde houve inserção/deleção de fronteiras

prosódicas). As células da coluna “palavras em períodos incongruentes”, por exemplo, indicam o número de palavras com hierarquia “x” encontradas em períodos incongruentes.

Tabela 51 – Classificação hierárquica das palavras em períodos congruentes e incongruentes (inserção/deleção de fronteiras)

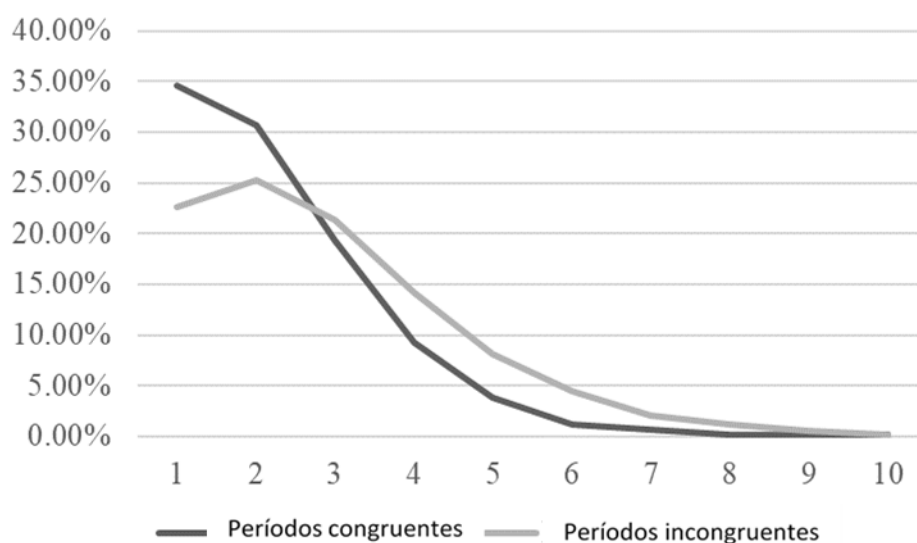
Hierarquia	Palavras em Períodos congruentes	%	Palavras em Períodos incongruentes	%
1	3852	34.5%	2088	22.5%
2	3413	30.6%	2334	25.2%
3	2148	19.3%	1973	21.3%
4	1037	9.3%	1316	14.2%
5	420	3.8%	755	8.1%
6	137	1.2%	413	4.5%
7	79	0.7%	192	2.1%
8	23	0.2%	105	1.1%
9	15	0.1%	53	0.6%
10	14	0.1%	21	0.2%
11	14	0.1%	6	0.1%
12	2	0.0%	4	0.0%
13	0	0.0%	6	0.1%
14	0	0.0%	5	0.1%
15	0	0.0%	5	0.1%
16	0	0.0%	1	0.0%
Total	11154	100.0%	9277	100.0%

Obs: ordenado pela coluna "Hierarquia"

Aplicando o teste do qui-quadrado (*Pearson's Chi-squared test with simulated p-value based on 2000 replicates*), concluímos que há algum tipo de associação entre a classificação hierárquica das palavras e a ocorrência de incongruências do tipo inserção/deleção de fronteiras, com um *p-value* de 0,0004998.

A **Figura 30** mostra o percentual de cada classificação hierárquica em períodos congruentes e incongruentes, sugerindo uma diferença na incidência das hierarquias de nível 1 e 2 em períodos congruentes.

Figura 30 – Percentual de cada posição hierárquica em períodos congruentes e incongruentes



Para identificar melhor essa associação entre classificação hierárquica e incongruência, agrupamos as 16 hierarquias da **Tabela 51** em apenas duas categorias, e de duas formas diferentes:

1. Hierarquia “1” vs. todas as demais hierarquias (2, 3, 4 etc)
2. Hierarquias “1” e “2” vs. todas as demais (3, 4, 5 etc)

Os resultados desses agrupamentos estão na **Tabela 52**.

Tabela 52 – Classificação hierárquica das palavras em períodos congruentes e incongruentes
(inserção/deleção de fronteiras)

Hierarquia	Período congruente	%	Período incongruente	%
1	3852	35%	2088	23%
não 1	7302	65%	7189	77%
1 + 2	7265	65%	4422	48%
não 1 + 2	3889	35%	4855	52%

Aplicando o teste do qui-quadrado (*Pearson's Chi-squared test with Yates' continuity correction*) aos dados da **Tabela 52**, observamos que continua existindo associação entre as distribuições hierárquicas das palavras e a ocorrência de incongruências, mesmo quando agrupamos

as classificações hierárquicas em apenas duas categorias. Além disso, observamos que o *p-value*, neste caso, sugere uma associação ainda mais significativa ($p\text{-value} < 2,2e-16$).

Para compreender melhor esse resultado, calculamos as contagens esperadas para cada categoria e as ilustramos na **Tabela 53**.

Tabela 53 – Classificação hierárquica das palavras em períodos congruentes e incongruentes:

ocorrências observadas/ocorrências esperadas

Hierarquia	Período congruente (observado/esperado)	Período incongruente (observado/esperado)
1	3852/3243	2088/2697
não 1	7302/7911	7189/6580
1 + 2	7265/6380	4422/5307
não 1 + 2	3889/4774	4855/3970

Obs: ocorrências esperadas calculadas de acordo com o *Pearson's Chi-squared test with Yates' continuity correction*

Esses resultados sugerem que tanto a classificação hierárquica 1 quanto a 2 estão associadas a períodos congruentes, enquanto as demais parecem associar-se a períodos incongruentes.

Como estamos analisando incongruências suprasegmentares e as inserções/deleções de fronteiras ocorrem entre palavras adjacentes, e não dentro de cada palavra, parece lógico estender a análise para a relação entre palavras adjacentes, em vez de analisar apenas a posição hierárquica isolada de cada palavra. Reconhecemos que os fenômenos prosódicos relacionados com as fronteiras, estudados do ponto de vista fonológico ou fonético, também ocorrem durante a produção dos fenômenos segmentares que dão origem à emissão sonora das palavras, mas isso não invalida a observação de que a inserção ou não inserção de fronteiras prosódicas são percebidas como fenômenos que segmentam ou dão continuidade a um conjunto de palavras emitidas em série, uma após a outra. Ou seja, separam ou aglutinam palavras.

Dessa forma, na frase hipotética “A maçã caiu do pé”, podemos analisar a classificação hierárquica de duas maneiras (entre outras tantas possíveis):

- a. Palavras isoladas: “A(1) maçã(1) caiu(1) do(2) pé(2)”
- b. Relação entre palavras adjacentes e sucessivas: “A(0-1)maçã(1-1)caiu(1-1)do(1-2)pé(2-2#)

Observe-se que empregamos as convenções arbitrárias “0” e “#”, respectivamente, em: “0-1” e “2-2#”, indicando que “A” é a primeira palavra do período, e que “pé” é a última palavra do período.

A **Tabela 54** mostra o número de ocorrências (eventos) de cada uma das relações hierárquicas observadas, em cada par de palavras adjacentes, em períodos incongruentes, e o número de inserções/deleções inadequadas de fronteira identificadas em cada uma dessas categorias de relação hierárquica.

Tabela 54 - Relação hierárquica entre palavras adjacentes em períodos incongruentes

Hierarquia	Eventos	Inserção/deleção de fronteira
1-1	1424	12
2-2	1377	4
3-3	1105	8
4-4	650	3
1-2	427	43
2-3	366	60
5-5	360	1
2-1	340	14
3-4	284	21
3-2	266	41
4-3	213	25
6-6	200	0
4-5	186	22
0-1	146	0
5-4	137	13
1-3	135	31
0-2	131	0
2-4	105	10
7-7	92	2
5-6	89	5
3-1	81	11
6-5	62	9
3-5	58	10
4-2	57	9
3-3#	55	0
8-8	53	0
4-4#	51	0
2-2#	44	0
5-3	36	6
6-7	36	4
7-6	35	0
4-1	34	4
1-4	30	4
5-5#	30	0
4-6	27	1
0-3	26	0
9-9	24	0
7-8	23	2
6-4	22	3
5-7	22	1

Obs: lista ordenada pela coluna “eventos”, contendo as 40 classificações mais frequentes de um total de 130 classificações observadas no corpus

O teste do qui-quadrado (*Pearson's Chi-squared test with simulated p-value, based on 2000 replicates*) mostrou que há associação entre as variáveis “hierarquia” (relação hierárquica entre duas palavras adjacentes) e “inserção/deleção de fronteira”, com um *p-value* = 0,0004998. Para identificar melhor essa associação, ilustramos na **Tabela 55** os percentuais de ocorrência dos eventos (número de ocorrências de cada relação hierárquica em períodos incongruentes) e das inserções/deleções de fronteiras (número de ocorrências de incongruências em cada categoria de relação hierárquica).

Tabela 55 - Relação hierárquica entre palavras: percentuais de ocorrência de incongruência prosódica

Hierarquia	Eventos	%	Ins/del fronteira	%
2-3	366	4%	60	14%
1-2	427	5%	43	10%
3-2	266	3%	41	10%
1-3	135	1%	31	7%
4-3	213	2%	25	6%
4-5	186	2%	22	5%
3-4	284	3%	21	5%
2-1	340	4%	14	3%
5-4	137	1%	13	3%
1-1	1424	15%	12	3%
3-1	81	1%	11	3%
2-4	105	1%	10	2%
3-5	58	1%	10	2%
6-5	62	1%	9	2%
4-2	57	1%	9	2%
3-3	1105	12%	8	2%
5-3	36	0%	6	1%
5-6	89	1%	5	1%
2-2	1377	15%	4	1%
6-7	36	0%	4	1%
4-1	34	0%	4	1%
1-4	30	0%	4	1%
4-4	650	7%	3	1%
6-4	22	0%	3	1%
7-7	92	1%	2	0%
7-8	23	0%	2	0%
5-5	360	4%	1	0%
4-6	27	0%	1	0%
5-7	22	0%	1	0%
0-x	328	4%	0	0%
6-6	200	2%	0	0%
3-3#	55	1%	0	0%
8-8	53	1%	0	0%
4-4#	51	1%	0	0%
2-2#	44	0%	0	0%
7-6	35	0%	0	0%
5-5#	30	0%	0	0%
9-9	24	0%	0	0%
6-6#	22	0%	0	0%
1-1#	21	0%	0	0%

a) lista ordenada pela coluna “inserção/deleção de fronteiras (%)”, contendo as 40 classificações mais frequentes de um total de 130 classificações

Pode-se observar que (pelo menos) as seis primeiras colocadas na **Tabela 55** são responsáveis por um percentual de inserção/deleção de fronteiras bem maior do que seu percentual de ocorrências no corpus. Além disso, observa-se, ao longo da lista, que há diversas relações hierárquicas (1-1, 2-2, 3-3...) com percentual de inserção/deleção de fronteiras bem menor do que seu respectivo percentual de ocorrências no corpus.

Para evidenciar melhor essas discrepâncias, reorganizamos a **Tabela 55**, classificando as relações hierárquicas em apenas dois grupos: relação hierárquica “nivelada”; p.ex. “1-1”, “2-2” etc; e relação hierárquica “desnivelada”; p.ex. “1-2”, “2-1”, “5-8” etc. Os resultados estão na tabela de contingência **Tabela 56**. Note-se que, por se tratar de uma tabela de contingência, para cada categoria de relação hierárquica (nivelada e desnivelada), foram discriminadas as ocorrências com e sem incongruência de fronteira. Em outras palavras: os eventos contados são transições entre palavras (relações hierárquicas) onde houve incongruência e onde não houve incongruência.

Tabela 56 - Tipo de relação hierárquica vs. inserção/deleção de fronteiras

	Congruentes	Incongruentes	Total
Nivelada	5510 (62%)	30 (7%)	5540
Desnivelada	3359 (38%)	378 (93%)	3737
<i>Total</i>	<i>8869 (100%)</i>	<i>408 (100%)</i>	<i>9277</i>

O teste do qui-quadrado (*Pearson's Chi-squared test with Yates' continuity correction*) aplicado à **Tabela 56** demonstra associação muito forte entre o tipo de relação hierárquica (nivelada vs. desnivelada) e congruência/incongruência de fronteira ($p\text{-value} < 2,2e-16$). Em outros termos, a incidência de inserção/deleção de fronteiras é muito menor quando a relação hierárquica é nivelada; ou é muito maior quando essa relação é desnivelada. Entre as relações niveladas de 1 a 4 (1-1, 2-2, 3-3 e 4-4), o teste do qui-quadrado (*Pearson's Chi-squared test with simulated p-value based on 2000 replicates*) não revela diferença significativa na incidência de incongruências de fronteira ($p\text{-value}=0,23$).

Se estendermos a análise da diferença entre relações niveladas e desniveladas aos períodos congruentes e incongruentes, em vez de apenas às incidências de incongruência em períodos incongruentes, teremos como resultado a **Tabela 57**, mostrando a quantidade de relações hierárquicas

de cada tipo (nivelada ou desnivelada) em cada tipo de período (congruente ou incongruente). O teste do qui-quadrado aplicado aos dados desta tabela (*Pearson's Chi-squared test with Yates' continuity correction*) também confirma a associação entre desnivelamento hierárquico e incongruência de períodos, com um *p-value* de 0,001512.

Tabela 57 - Tipo de relação hierárquica vs. congruência e incongruência de períodos

	Congruentes (observados/esperados)	Incongruentes (observados/esperados)
Nivelada	6902/6792	5437/5547
Desnivelada	4182/4293	3614/3504

É necessário analisar, também, a relação entre desnivelamento hierárquico e cada tipo específico de inserção/deleção incongruente de fronteira. Nessa análise, deixamos de fora a categoria “deleção de fronteira terminal”, pois ela teve relativamente poucos eventos, prejudicando a força estatística da análise. Para fazer a análise, segregamos os dados por tipo de incongruência (INT: inserção de fronteira não terminal; IT: inserção de fronteira terminal; DNT: deleção de fronteira não terminal) e aplicamos o teste do qui-quadrado para cada incongruência (“hierarquia” vs. INT; “hierarquia” vs. IT; “hierarquia” vs. DNT). Os resultados estão descritos na **Tabela 58**. Essa tabela mostra cada tipo de incongruência de fronteira e as relações hierárquicas mais comuns em que essas incongruências ocorreram. Os dados estão ordenados pela coluna “%”, que calcula “número de incongruências” (coluna “incongruências”) dividido por “número de eventos” (coluna “eventos”) para cada relação hierárquica (p.ex. “1-3”; “3-5” etc). A análise estatística confirmou uma associação entre relação hierárquica e cada subtipo de incongruência de fronteiras, com *p-values* de 0,0005; 0,0015; e 0,014 para, respectivamente, INT, IT e DNT.

Tabela 58 - Relações hierárquicas com maior incidência relativa de incongruências de inserção/deleção de fronteira

Hierarquia	Eventos	Incongruências	%
INT			
1-3	135	26	19%
3-5	58	8	14%
2-3	366	41	11%
4-5	186	18	10%
2-4	105	8	8%
IT			
3-1	81	8	10%
4-2	57	4	7%
3-2	266	15	6%
6-5	62	3	5%
4-3	213	10	5%
DNT			
3-2	266	11	4%
3-1	81	3	4%
4-2	57	2	4%
6-5	62	2	3%
2-1	340	10	3%

Obs: "eventos" = número de ocorrências da respectiva relação hierárquica; "incongruências" = número de incongruências de cada tipo (INT, IT, DNT) ocorridas na respectiva relação hierárquica; "%" = percentual da respectiva relação hierárquica em que ocorreram incongruências de cada tipo

Já havíamos observado que as relações hierárquicas niveladas (1-1, 2-2...) estão associadas a uma incidência bem menor de incongruências de fronteira. Nesta **Tabela 58**, notamos, adicionalmente, que as relações hierárquicas mais frequentemente associadas às incongruências do tipo INT são descendentes, ou seja, são relações desniveladas entre uma palavra de hierarquia mais alta e outra, subsequente, de hierarquia mais baixa (p.ex. 2-3). Em contrapartida, no caso das incongruências do tipo IT e DNT, as relações mais frequentes são ascendentes (p.ex. 3-2).

Para analisar e compreender melhor essa observação, agrupamos os dados da **Tabela 55** (incluindo os dados restantes, totalizando 130 tipos diferentes de hierarquia) em apenas duas categorias hierárquicas (ascendente e descendente), e confrontamos esses dados com os subtipos de incongruência. A **Tabela 59** mostra os resultados desse agrupamento, contando o número de vezes em que cada tipo de incongruência ocorreu, em associação com cada tipo de relação hierárquica (ascendente ou descendente), resultando em três tabelas de contingência.

Tabela 59 - Relações hierárquicas: "desnívelamento" vs. congruência

Desnívelamento	INT	Congruente
asc	45	1402
desc	153	1804
Desnívelamento	IT	Congruente
asc	61	1387
desc	47	1912
Desnívelamento	DNT	Congruente
asc	42	1406
desc	23	1936

Obs: "asc" = ascendente; "desc" = descendente

O teste do qui-quadrado (*Pearson's Chi-squared test with Yates' continuity correction*) aplicado a cada uma das subtabelas confirma a associação entre desnívelamento descendente e INT ($p\text{-value} = 1,018\text{e-}08$), bem como entre desnívelamento ascendente e IT ($p\text{-value} = 0,003878$) e desnívelamento ascendente e DNT ($p\text{-value} = 0,0004399$).

Outra análise que fizemos diz respeito à posição da inserção/deleção de fronteira dentro do período, independentemente da classificação hierárquica de suas palavras. Esse dado foi quantificado considerando a posição inicial do período como 0, e sua posição final como 1. Por exemplo, na frase hipotética “A tartaruga atravessou a pista / de pouso.” (onde “/” indica a inserção de uma fronteira não terminal), a posição da incongruência seria 5/7; ou seja, ocorreu depois da quinta palavra em um período de sete palavras. O resultado numérico computado foi decimal: $5/7 = 0,714285714$.

Os resultados foram coletados para INT, IT e DNT. A **Tabela 60** mostra as médias, medianas e o $p\text{-value}$ do teste de normalidade de Shapiro-Wilk para cada categoria de incongruência, além dos valores mínimo e máximo **observados** (note-se que não se trata de “intervalo de confiança”).

Tabela 60 - Incongruência: posição no período

Incongruência	Média	Mediana	Shapiro-Wilk ($p\text{-value}$)	Mín-Máx
INT	0,47	0,45	0,00008863	0,04-0,94
IT	0,59	0,62	0,007054	0,17-0,96
DNT	0,51	0,50	0,05949	0,02-0,92

Observa-se que, de acordo com o teste de Shapiro-Wilk, a distribuição dos valores de posição das incongruências nos períodos não segue uma distribuição normal. Isso significa que não há uma posição específica no período que tenha a propriedade de atrair determinado tipo de incongruência. Na

verdade, as incongruências ocorreram ao longo de todo o período, do início ao fim, como se pode observar na coluna “Mín-Máx” da tabela.

Entretanto, o valor mínimo para IT foi superior a INT e DNT, e suas média e mediana também foram mais altas. Para avaliar se a distribuição de IT realmente tende a concentrar-se em posições mais tardias nos períodos, aplicamos o teste de Kruskal-Wallis (*Kruskal-Wallis rank sum test*) aos dados de posição no período de cada tipo de incongruência e obtivemos os resultados descritos na **Tabela 61**:

Tabela 61 - Incongruências: comparando as posições no período

Incongruências	Kruskal-Wallis <i>p-value</i>
INT vs. IT vs. DNT	0,0000664
INT vs. DNT	0,2812
INT vs. IT	0,000009516
IT vs. DNT	0,03452

Confirma-se que as posições de IT realmente são discrepantes com relação a INT e DNT, ocorrendo mais perto do final do período, ou, alternativamente, menos perto do início do período. Tal resultado seria intuitivamente esperado, por tratar-se de fronteiras terminais, mas uma interpretação menos superficial do fenômeno será feita, tentativamente, no capítulo Discussão.

Analizamos também a posição da incongruência com relação à predicação primária do período. Considerando como exemplo o período hipotético “No domingo retrasado, **uma tartaruga atravessou lentamente a pista** de pouso e foi vista por um piloto.”, o trecho em negrito corresponderia à predicação primária: “uma tartaruga atravessou a pista”. Nesse período, as incongruências podem ocorrer antes, durante ou depois da predicação primária, como exemplificado abaixo (utilizando a notação “/”, correspondente a INT):

- Antes: No domingo / retrasado, **uma tartaruga atravessou lentamente a pista** de pouso e foi vista por um piloto.
- Durante: No domingo retrasado, **uma tartaruga atravessou lentamente / a pista** de pouso e foi vista por um piloto.
- Depois: No domingo retrasado, **uma tartaruga atravessou lentamente a pista / de** pouso e foi vista por um piloto.

A **Tabela 62** mostra o número de ocorrências de cada tipo de incongruências em cada posição relativa à predicação primária.

Tabela 62 - Incongruências: posição relativa à predicação primária

	INT	%	IT	%	DNT	%
Antes	23	13%	13	15%	14	32%
Durante	37	21%	6	7%	9	20%
Depois	118	66%	70	79%	21	48%

A **tabela** sugere, ao menos, duas diferenças entre as incongruências. As inserções de fronteira (INT e IT) parecem ter uma incidência maior “depois” da predicação primária; e nas IT, essa discrepância parece ser ainda mais intensa. O teste do qui-quadrado (*Pearson's Chi-squared test*) confirma a significância estatística dessas diferenças, com *p-values*, respectivamente, de 0,003663 (INT+IT vs. DNT) e 0,01304 (INT vs. IT).

4.4.1 Resumo dos resultados desta seção:

- Descrições
 - IT costuma ocorrer mais perto do final do período, em comparação com INT e DNT
 - IT e INT ocorrem, preferencialmente, depois da predicação primária, sendo que a preferência é mais acentuada para IT
- Análises
 - Associam-se com congruência
 - Classificação hierárquica 1 ou 2
 - Relação hierárquica nivelada
 - Associam-se com incongruência (de inserção/deleção de fronteiras)
 - Classificação hierárquica maior do que 1 ou 2
 - Relação hierárquica desnivelada
 - Se descendente: INT
 - Se ascendente: IT e DNT

5. DISCUSSÃO

Nosso trabalho pretendeu estudar um corpus de fala lida, identificando as incongruências produzidas pelos leitores e buscando correlações entre essas incongruências e determinados aspectos linguísticos dos textos escritos. Nossas hipóteses iniciais eram as seguintes:

- determinados aspectos morfossintáticos, semânticos e/ou discursivos dos textos escritos poderiam estar associados a maior ou menor frequência de incongruências de leitura
- tais associações poderiam refletir aspectos fundamentais de mecanismos de processamento da linguagem

Para testar essas hipóteses, lemos textos escritos e assistimos a vídeos/áudios das leituras realizadas, identificando as incongruências produzidas. Na sequência, voltamos aos textos escritos e identificamos alguns aspectos linguísticos previamente escolhidos. As variáveis obtidas (incongruências e aspectos linguísticos, além de variáveis demográficas) foram analisadas e correlacionadas estatisticamente. Os resultados descritos no capítulo anterior atenderam aos três primeiros objetivos específicos mencionados na Introdução, quais sejam:

- identificar, sistematicamente, as incongruências mais comuns em nosso corpus de falas lidas (pronunciamentos parlamentares)
- identificar, nos textos escritos que deram origem às falas lidas, determinadas características linguísticas (morfossintáticas, semânticas e discursivas) pré-selecionadas
- analisar as correlações estatísticas eventualmente existentes entre incongruências e características linguísticas

Resta-nos, então, atender aos demais objetivos específicos, quais sejam:

- analisar as correlações estatísticas à luz de teorias atuais de processamento da linguagem
- discutir se os achados analisados poderiam, eventualmente, contribuir para o aperfeiçoamento de teorias atuais de processamento de sentenças, ou teorias cognitivas de um modo geral

- elaborar, com base nos achados do estudo, possíveis recomendações para aperfeiçoar a produção de textos escritos cujo objetivo seja a leitura e voz alta, especialmente quando voltadas ao fonostilo político

Para isso, organizamos este capítulo, inicialmente, como um espelho do capítulo anterior (Resultados), discutindo cada resultado em uma ordem semelhante à ordem em que foi descrito e (parcialmente) analisado. Após essa discussão inicial, acrescentamos algumas discussões adicionais necessárias ao atendimento dos objetivos, e alguns tópicos que julgamos merecedores de uma atenção especial.

5.1 Os falantes, o corpus e as incongruências *lato sensu*

5.1.1 Os falantes

A amostra de falantes estudada em nosso corpus, do ponto de vista da diversidade e da representatividade com relação a falantes de PB (a população brasileira), tem pontos fortes, pontos fracos e algumas peculiaridades.

Como pontos fortes, podemos mencionar a grande diversidade diatópica (de origem geográfica): são falantes nascidos em 20 estados brasileiros diferentes, representando todas as grandes regiões do País. Apesar de não termos considerado a naturalidade do falante como uma das variáveis do estudo — e não o fizemos porque temos poucos representantes de cada estado —, acreditamos que, no conjunto, a pluralidade seja um ponto positivo.

Outro aspecto da amostra de falantes que pode ser considerado, ao mesmo tempo, positivo e negativo, é a distribuição de idades. Trata-se de uma população com idade média em torno de 63 anos, com desvio padrão de 11 anos, com poucos representantes entre 39 e 50 anos e uma boa quantidade acima de 70 anos. A população brasileira, segundo projeções do IBGE, é bem mais jovem, com quase 90% de indivíduos abaixo de 63 anos. A idade média relativamente avançada de nossa amostra pode ser considerada positiva por representar uma faixa etária que, ao que nos parece, não costuma ser habitualmente representada em estudos linguísticos. Dessa forma, estudos que abordem, em falantes mais jovens, quaisquer das variáveis que estudamos, podem trazer informações relevantes quanto a possíveis variações linguísticas relacionadas ao envelhecimento.

O envelhecimento, por sua vez, pode constituir um aspecto negativo de nossa amostra. Nosso estudo é retrospectivo e, por definição, possui certas limitações. Não esteve ao nosso alcance, por exemplo, o controle de variáveis como a acuidade visual dos falantes ou a formatação dos textos utilizados para a leitura (tamanho de fonte, espaçamento, quebras de linha etc). Por se tratar de falantes com mais idade, e de uma tarefa de leitura, é possível que exista influência da acuidade visual na ocorrência de incongruências. O que pudemos fazer, e fizemos, com relação a essa possibilidade, foi testar a influência da idade (considerando as idades presentes em nossa amostra de falantes) sobre a incidência de incongruências.

Ainda sobre os falantes, pode-se observar que não há grande variação diastrática (social). Tomando a liberdade de inferir a posição social pelo grau de instrução — e, obviamente, pela função ocupada pelos falantes —, notamos que apenas 3 deles não têm curso superior. Esse aspecto pode ser considerado negativo, por diminuir a representatividade da amostra com relação ao universo de falantes do PB. Por outro lado, temos uma amostra mais homogênea e mais representativa de um estrato social específico, o que significa que, em virtude dessa homogeneidade, nossos resultados têm mais força estatística com relação ao estrato social que compõe a amostra: falantes com nível superior, socialmente proeminentes, e de idade mais avançada.

Outro aspecto negativo de nossa amostra é o desequilíbrio entre os sexos. Temos apenas 7 mulheres entre nossos falantes (menos de um quinto da amostra), o que desloca nossos resultados para uma realidade predominantemente masculina, e isso pode não representar a realidade linguística da população geral, onde a regra é o equilíbrio.

Apesar dessas observações, consideramos que, modo geral, nossa amostra de falantes seja adequada para os fins propostos pelo estudo.

5.1.2 O corpus

O corpus analisado foi composto por 5 horas de falas lidas, computada a soma dos tempos de duração dos períodos. Trata-se de um tempo total de falas lidas razoável para análise, com uma média de 8 minutos de leitura por falante. Deve-se notar, entretanto, a grande variabilidade nos tempos de leitura individual de cada falante, estendendo-se de 1 a 19 minutos. A "oportunidade" de cometer incongruências, portanto, não está distribuída uniformemente entre os falantes, o que nos obriga a considerar que nossos resultados podem ter um viés em prol daqueles indivíduos que produziram mais tempo de leitura. Essa observação, entretanto, não causa grande prejuízo ao estudo, uma vez que não

estamos investigando a influência de características demográficas específicas sobre a língua, mas sim, a influência de características linguísticas do texto escrito sobre o desempenho de um falante genérico, um representante abstrato dos falantes de PB, que é representado, em nosso trabalho, pelo conjunto de falantes.

5.1.3 As incongruências

5.1.3.1 Os períodos congruentes e incongruentes

Todos os períodos lidos foram classificados em congruentes ou incongruentes e, nos períodos incongruentes, identificamos e classificamos cada incongruência encontrada. O primeiro resultado que nos surpreendeu foi que todos os falantes produziram períodos incongruentes durante a leitura. Não houve, entre os 39 falantes participantes do estudo, nenhum que tenha feito sua leitura sem produzir algum tipo de incongruência. Dentre os 39 falantes, 29 produziram mais períodos com incongruências do que sem incongruências; e, no cômputo geral, tivemos um total de 854 períodos incongruentes e de apenas 575 sem nenhuma incongruência. Esse resultado, isoladamente, nos revela que a congruência, se conceituada, analogamente à fluência, como "uma combinação de acurácia, automaticidade e (...) prosódia que, tomados em conjunto, facilitam a reconstrução do sentido pelo leitor; e que se demonstra, na leitura oral, por meio da facilidade de reconhecer as palavras e de imprimir ritmo, fraseamento e entoação adequados [ao texto]" (KUHN, 2010, p. 240), não é alcançada de maneira trivial na fala lida. Ou, para ser mais específico, não o é no português brasileiro (ou no 'português padrão brasileiro'), ou, pelo menos, não o foi em nosso corpus.

5.1.3.2 Congruência vs. idade

Observamos que a produção de incongruências não tem correlação com a idade. Leitores mais velhos não produziram menos incongruências (o que se poderia justificar pela maior experiência) ou mais incongruências (o que se explicaria pelo envelhecimento) do que os mais jovens. Novamente, reconhecemos as limitações de um estudo retrospectivo e lembramos que um estudo experimental, com controle mais efetivo das variáveis independentes e dos fatores de confusão, nos traria resultados mais confiáveis. Essa observação quanto ao desenho do estudo, na verdade, aplica-se a todos os resultados de nosso trabalho, e não será, via de regra, retomada nesta discussão.

5.1.3.3 Congruência vs. sexo

No que se refere a outra variável demográfica — o sexo —, houve correlação entre sexo feminino e congruência. As sete mulheres do estudo produziram 46% de períodos incongruentes, enquanto os 32 homens produziram 65%, e tal diferença é estatisticamente significativa. A literatura não costuma estudar a congruência de leitura em indivíduos adultos, dando ênfase a crianças e jovens em idade escolar, o que sugere que as incongruências de leitura são encaradas como um fenômeno transitório, que se resolve quando se completa um ciclo de aprendizado e já se adquiriu experiência suficiente com a escrita e a leitura. Na faixa etária que se costuma estudar, há reais evidências de que as meninas e moças são mais desenvoltas nas tarefas de leitura do que os meninos e rapazes, como demonstra, entre outros, o estudo de Torppa, Eklund, Sulkunen, Niemi e Ahonen (2017).

Esses autores partiram dos resultados do PISA (Programa Internacional de Avaliação de Estudantes) — que avalia as habilidades e conhecimentos em matemática, ciências e leitura em jovens de 15 anos de idade ao redor do mundo — para tentar explicar porque as moças têm resultados consistente e onipresentemente melhores do que os rapazes nos testes de leitura. Eles testaram 1309 jovens finlandeses participantes do PISA e os submeteram a alguns experimentos, incluindo testes de congruência de leitura. Esses testes consistiram em três tarefas: ler o máximo possível de períodos curtos em 2 minutos e marcar certo ou errado (p.ex.: "bolas são redondas"); encontrar e marcar o máximo possível de erros ortográficos em um texto, em 3 minutos; e indicar o local de segmentação entre palavras que haviam sido escritas sem espaços em branco entre elas, em 90 segundos. Diferentemente do nosso estudo, os testes de congruência empregados não incluíram uma avaliação do desempenho na leitura em voz alta, mas nos parece que, de certa forma, existe paralelo entre o desempenho nas tarefas do estudo e a capacidade de ler em voz alta sem produzir incongruências, tanto segmentares (identificação correta de palavras gráficas), quanto suprasegmentares (compreensão rápida da estrutura prosódica correta do período escrito). Os autores mostraram que os resultados de leitura do PISA, que favorecem as meninas, podem ser explicados por diferenças entre os sexos em comportamentos como leitura por prazer, realização de dever de casa, e tendência a superar desafios. Mas a diferença de desempenho de leitura entre moças e rapazes no PISA correlaciona-se melhor com o desempenho das moças nos testes de congruência aplicados no estudo. Podemos, portanto, com os riscos inerentes à empreitada, extrapolar esses resultados para justificar a diferença de desempenho entre as senhoras e os senhores participantes do nosso estudo e acrescentar que, talvez, as diferenças observadas na infância e na juventude sejam persistentes e se prolonguem por toda a vida. E, para tentar explicar as possíveis causas dessa diferença de congruência de leitura entre homens e mulheres,

poderíamos apelar para o antigo dilema do "*nature versus nurture*"¹. Do ponto de vista social, pode ser que as vantagens adquiridas até a adolescência sejam tão definitivas que não possam mais ser superadas. Ou, do ponto de vista natural, pode ser que as mulheres sejam geneticamente mais bem dotadas para a congruência de leitura do que os homens. Obviamente, não conseguiremos resolver esse dilema com nosso estudo.

5.1.3.4 Congruência vs. extensão dos períodos (e comentários adicionais)

Analizamos também a influência da extensão dos períodos (número de palavras por período) sobre a ocorrência de incongruências. Concluimos que, considerando a extensão habitual dos períodos em nosso corpus (entre 1 e 50 palavras, com poucos períodos acima dessa extensão), períodos maiores contêm mais incongruências. Esse resultado é esperado, uma vez que períodos maiores oferecem mais oportunidades de cometer incongruências. Se existe uma chance de incongruência em cada sílaba fonológica, em cada frase, em cada constituinte produzido na leitura, quanto mais sílabas, frases ou constituintes produzidos, maior a incidência de incongruência (soma das chances individuais).

O que nos chamou atenção foi que, ao contrário do que imaginávamos, a correlação entre extensão do período e ocorrência de incongruências foi linear. Isso significa que períodos considerados longos em relação à média de extensão dos períodos do corpus não têm uma chance desproporcionalmente maior de conter incongruências. Matematicamente falando, se dividirmos as chances de ocorrência de incongruências em períodos com 40 e com 20 palavras, essa divisão terá o mesmo resultado que a divisão das chances de ocorrência de incongruência entre períodos com 10 e com 5 palavras. Essa linearidade se aplica tanto às incongruências segmentares — o que faz bastante sentido, já que tais incongruências ocorrem dentro das palavras —, quanto às incongruências suprasegmentares, onde esperávamos observar um efeito desproporcional da extensão do período sobre a capacidade do falante de produzir uma estrutura prosódica coerente.

Em um estudo de fala espontânea, Oviatt (1995) observou uma relação não linear entre extensão do período e incidência de incongruências. Quanto maior o período, maior a incidência de incongruências, variando de 1 incongruência a cada 100 palavras em períodos com extensão de 5

¹ Simplificadamente, a expressão representa a dificuldade em separar o que é consequência da “natureza” (o que é genético, o que já está conosco quando nascemos, o que é ‘interno’), do que é consequência do ambiente em que um organismo se desenvolve (cultura, alimentação, e outros fatores ‘externos’)

palavras, a 4 incongruências a cada 100 palavras em períodos com extensão de 15 palavras. Entretanto, seu material de estudo era diferente do nosso: avaliou fala espontânea, interação com um sistema informatizado e períodos pequenos. Apesar das diferenças, é interessante notar que a incidência de incongruências em nosso corpus de leitura (6,6 incongruências por 100 palavras) é semelhante à incidência de incongruências nos diferentes materiais avaliados no estudo de Oviatt (1995): 8,83 em conversa telefônica, 5,50 em conversa presencial, 3,60 em um monólogo, e 1,80 na interação com um computador.

Não podemos esquecer, entretanto, que aquele estudo não computou incongruências suprasegmentares de fronteira, porque, ao que parece, elas não são características da fala espontânea. Considerando, em nosso estudo, apenas as pausas e as hesitações (estas últimas denominadas, por eles, de pausas preenchidas), e descontando as outras incongruências suprasegmentares, tivemos, ainda assim, 5,13 incongruências por 100 palavras, o que mantém a fala lida analisada em nosso corpus alinhada com a fala espontânea do estudo citado com relação à incidência de incongruências.

Nesta altura, devemos ressaltar um fenômeno que, a julgar por sua irrelevância na literatura que avaliamos, parece diferenciar a fala espontânea da fala lida: a presença de incongruências suprasegmentares de fronteira na fala lida. Um estudo bastante citado sobre incongruências em fala espontânea (BORTFELD; LEON; BLOOM; SCHOBBER; BRENNAN, 2001) parece confirmar a observação de que as incongruências de fronteira não chamam atenção na fala espontânea. Nesse estudo, elas não foram incluídas entre as variáveis analisadas em um corpus de 192 mil palavras pronunciadas por 48 pares de indivíduos, em situação de conversação espontânea baseada em assuntos pré-determinados. As incongruências computadas foram repetições de palavras (*just on the left left side*)¹, hesitações e recomeços (*restarts*, como em *imme- just below the left side*)²; e o que se observou foi que: a) não houve aumento significativo da incidência de incongruências com a idade; b) tarefas com maior exigência cognitiva estão associadas a maior incidência de incongruências; e c) as hesitações devem estar mais associadas à organização dos turnos discursivos do que propriamente a incongruências. O item "a" concorda com nossos achados. O item "b", nós não avaliamos, até porque

¹ Bem do lado esquerdo esquerdo. (tradução nossa)

² Imme- bem debaixo do lado esquerdo. (tradução nossa)

o que se considera cognitivamente exigente em fala espontânea não encontra paralelo na fala lida. E, finalmente, o item "c" não se aplica a monólogos, nem à leitura.

A linearidade na incidência de incongruências de acordo com a extensão dos períodos observada em nosso estudo precisa ser explicada. Como vimos, as incongruências suprasegmentares mais frequentes são as pausas e as inserções de fronteira não terminal, que compõem 62% de todas as incongruências suprasegmentares (534 eventos, num total de 858 incongruências suprasegmentares *lato sensu*). As pausas constituem interrupções aleatórias do fluxo da fala, e parecem ocorrer quando o falante precisa de mais tempo para processar algo que ele não conseguiu processar na velocidade (taxa de elocução) de leitura atual. Portanto, a extensão dos períodos, por si só, não deve implicar maior complexidade de processamento, a ponto de aumentar a chance de inserção de pausas durante a leitura em voz alta. Devemos lembrar que a leitura em voz alta não exige compreensão completa do que está sendo lido, como já mencionamos (DEHAENE, 2009). É possível que um experimento que envolva uma tarefa de compreensão durante a leitura em voz alta consiga estabelecer melhor a relação entre extensão de períodos e inserção de pausas na leitura, podendo até mesmo revelar uma correlação não linear, pela maior demanda sobre a memória em períodos muito extensos.

A segunda incongruência suprasegmentar mais frequente, a inserção de fronteira não terminal, é diferente da pausa por estar na categoria de incongruências de fronteira. Trata-se, aqui, de uma incongruência que indica uma dificuldade para apreender e reproduzir uma estrutura prosódica que, por sua vez, reflete a organização sintática e semântica do período. Ela é inserida para delimitar determinados segmentos de um período, e esse raciocínio pode ser extrapolado para concluirmos que a) períodos maiores terão mais segmentos e, portanto, mais oportunidades de inserção de fronteiras não terminais; mas que b) o aumento de oportunidades, de segmentos reais ou imaginários, é diretamente proporcional ao aumento de palavras no período, crescendo linearmente com o crescimento do período. Esta seria uma possível justificativa para a linearidade observada na relação entre extensão dos períodos e ocorrência de incongruências suprasegmentares: numerador (inserção de fronteira não terminal) e denominador (número de palavras) cresceriam no mesmo ritmo.

5.1.3.5 Congruência vs. taxas de elocução

Outra variável analisada foi a taxa de elocução, medida em palavras gráficas lidas por minuto. Considerando os períodos lidos congruentemente como referência para a taxa de elocução, calculamos que os participantes de nosso estudo leem com uma "velocidade" média de 120 palavras por minuto,

ou duas palavras por segundo. Essa taxa, entretanto, varia bastante entre indivíduos, indo de 82 a 179 palavras por minuto, com grande concentração de falantes entre 90 e 140 palavras por minuto. Para analisar a relação dessa variável com a ocorrência de incongruências, partindo do pressuposto de que taxas de elocução maiores poderiam estar associadas a maior ocorrência de incongruências (assim como maior velocidade no trânsito associa-se a mais e mais graves acidentes), comparamos as taxas de elocução dos períodos congruentes com as dos períodos incongruentes. Observamos que as taxas de elocução são realmente diferentes entre essas duas categorias de períodos, mas as maiores taxas de elocução (maior velocidade) ocorreram nos períodos congruentes, ao contrário do que imaginávamos. Percebemos, então, que, enquanto as taxas de elocução em períodos congruentes são reais (computam apenas palavras lidas; sem erros, retratações ou interrupções involuntárias), as taxas de elocução nos períodos incongruentes excluem do numerador incongruências e retratações, e incluem no denominador interrupções indevidas do fluxo da fala (maior tempo de fala). Em função disso, as taxas de elocução em períodos incongruentes não correspondem às taxas de elocução reais, e precisaríamos de uma série de adaptações para torna-las uma medida útil para nossas análises.

Por isso, resolvemos trabalhar apenas com a taxa de elocução em períodos congruentes, acreditando que ela corresponderia à taxa de elocução média real de cada falante; ou seja, à maneira habitual com que cada falante abordaria o texto escrito ao realizar uma leitura em voz alta. E, partindo dessa variável, correlacionamo-la com o percentual de períodos incongruentes produzidos por cada falante. Como vimos, não houve associação entre essas variáveis, indicando que indivíduos que optam por ler mais rápido em voz alta não produzem mais incongruências do que os que optam por ler mais devagar. Tal achado suscita algumas discussões adicionais.

Em primeiro lugar, é possível que a taxa de elocução seja escolhida voluntariamente por cada leitor na medida de sua própria capacidade de ler sem produzir incongruências. Sob essa perspectiva, o que limitaria a taxa de elocução de cada falante seria a ocorrência de incongruências e, assim, cada indivíduo restringiria sua velocidade para não produzir mais incongruências. Uma segunda possibilidade é que haja uma compensação entre velocidade e congruência, sendo que velocidades maiores poderiam, ao mesmo tempo, dificultar e facilitar a leitura. Taxas de elocução maiores dificultariam a leitura por razões óbvias, de articulação e de processamento da linguagem; e a facilitariam por permitirem uma visão mais ampla do período, reduzindo as incongruências suprasegmentares e, como veremos adiante, demandando menos movimentos respiratórios dentro de cada período. Novamente, esbarramos nas limitações de um estudo retrospectivo, não experimental. É bem provável que um experimento que manipule apenas a variável "taxa de elocução", mantendo as

demais variáveis inalteradas, identifique que taxas de elocução maiores, via de regra, estão associadas a maior incidência de incongruências, de qualquer tipo.

Além dessas observações, vale a pena retomar uma discussão iniciada no capítulo Metodologia, justificando o uso de palavras gráficas como o numerador da taxa de elocução, em lugar de sílabas fonológicas ou unidades v-v. É fenômeno conhecido, como mencionamos, que diferentes línguas têm número médio de sílabas fonológicas por palavra diferentes (BARBOSA, 2019). Portanto, uma taxa de elocução medida em sílabas ou unidades v-v por tempo representaria muito bem a velocidade de produção sonora, mas não tão bem a velocidade de produção (e processamento) de informação. Línguas com maior número de sílabas por palavra terão maior velocidade de produção sonora, mas, não necessariamente, maior velocidade de produção e transmissão de informação.

Pimentel, Roark e Cotterell (2020) demonstram que a complexidade fonotática de uma língua é inversamente proporcional ao tamanho médio de suas palavras. Fonotaxes mais complexas, com maior número de possibilidades de combinação entre fonemas, resultam em maior carga de informação por fonema e, conseqüentemente, em palavras menores, o que manteria a "taxa de informação" (*information rate* — quantidade de informação transmitida por unidade de tempo) comparável entre as diversas línguas. E Coupé, Oh, Dediu e Pellegrino (2019) compararam 17 línguas (infelizmente, não incluíram o português) e observaram uma compensação entre taxa de elocução (em sílabas fonológicas por unidade de tempo, *speech rate*) e a quantidade de informação por sílaba. Línguas em que os falantes produzem mais sílabas por segundo (como japonês e espanhol, com cerca de 8 sílabas/s) transmitem a mesma quantidade de informação (em bits por segundo, cerca de 39 bits/s) do que línguas que produzem menos sílabas por segundo (como o vietnamita e o tailandês, com cerca de 5 sílabas/s).

Portanto, para os fins a que se destina o nosso trabalho, o uso de palavras por unidade de tempo como indicativo da taxa de elocução estaria mais diretamente relacionado, para fins comparativos com outras línguas, ao processamento de informações e a fenômenos cognitivos. Ou seja, partimos do ponto de vista de que as palavras são unidades de informação autônomas das línguas e, assim, são unidades mais comparáveis do que sílabas ou unidades v-v, cuja utilidade deve estar mais restrita a cada língua em particular.

5.1.3.6 Os padrões respiratórios

Nosso estudo mostrou que, em média, as respirações na leitura ocorrem a cada 2,8 segundos, com um desvio padrão de meio segundo; e que são emitidas, entre cada respiração, pouco mais de 5

palavras. Entretanto, precisamos considerar algumas características metodológicas que influenciaram esses valores.

Em primeiro lugar, nas situações (não tão comuns) em que houve uma pausa muito longa entre dois períodos adjacentes, o falante respirou duas vezes: logo após terminar um período; e imediatamente antes de iniciar o próximo. Nesses casos, registramos uma respiração a mais sem que tenha havido nenhuma emissão sonora antes da próxima inspiração. Isso fez com que o cálculo de palavras entre respirações tenha tido um resultado um pouco menor do que a realidade, pois o denominador da divisão (número de respirações) foi um pouco maior.

Além disso, nosso valor de 2,8 segundos entre respirações incluiu todo o tempo respiratório (incluindo o tempo de inspiração, sem fonação), pois não tínhamos como avaliar a extensão do movimento inspiratório e descontá-lo do cálculo. Considerando isso, acreditamos que o tempo efetivo entre o final de uma inspiração e o início da próxima inspiração seja um pouco menor do que o que encontramos. De qualquer forma, como não nos propusemos a medir exclusivamente o tempo expiratório, acreditamos que a variável "tempo entre respirações", tendo sido definida como foi, é válida e traz informações relevantes. Além disso, considerando as imperfeições metodológicas relativas à avaliação dos períodos incongruentes, já explicadas no capítulo Resultados, preferimos considerar, para efeito de comparação com dados da literatura, os valores obtidos, exclusivamente, a partir dos períodos congruentes: 3,33 segundos entre respirações; e 6,54 palavras entre respirações.

O tempo efetivo entre o final de uma inspiração (que coincide com o início da expiração) e o início da próxima inspiração (que coincide com o final da expiração e, conseqüentemente, da produção sonora) é chamado, por alguns, de *breath group* (BG). Estudos com um grupo de 6 jovens do sexo feminino, falantes nativas de inglês, em tarefas de fala espontânea (WINKWORTH; DAVIS; ADAMS; ELLIS, 1995) e fala lida (WINKWORTH; DAVIS; ADAMS; ELLIS, 1994) mostram diferenças entre os padrões respiratórios nessas duas situações. Nesses estudos, a respiração foi aferida por meio da técnica de pletismografia respiratória por indutância, que mede os movimentos respiratórios do tórax e abdome por meio de faixas elásticas transdutoras, e captura os movimentos e os volumes inspiratórios e expiratórios. No estudo com fala espontânea, observou-se um BG médio de 3,8 segundos, com desvio padrão de 2 segundos. E no estudo com fala lida, o BG (chamado *utterance length*) médio foi de 3,36 segundos, com desvio padrão de 1,39.

Perceba-se que não podemos comparar os desvios padrão encontrados nesses estudos com o desvio padrão do nosso estudo, pois utilizamos as médias de cada falante, enquanto os autores utilizaram todos os eventos do estudo, independentemente do falante. Nosso desvio padrão, portanto, indica diferenças entre falantes, não entre BGs. Mas nosso achado de um BG de 3,33 segundos para os períodos de leitura congruente pode ser comparado, guardadas as diferenças metodológicas, com o valor de 3,36 encontrado pelos autores para a fala lida.

Os autores também concluem que a variabilidade de extensão dos BGs é bem menor na fala lida do que na fala espontânea. E acreditam que a menor variabilidade na fala lida estaria representada na literatura em estudos que não revelam diferenças significativas de BGs (na fala lida) entre sexos, entre pessoas de idades diferentes, ou mesmo entre pessoas saudáveis e pessoas com asma brônquica, por exemplo. Esses dados são compatíveis, novamente, com nossos resultados, que não mostraram correlação entre idade/sexo e tempo entre respirações.

A estabilidade dos BGs em tarefas de leitura, a despeito das diferenças físicas e fisiológicas inegáveis que existem entre idades e sexos diferentes, estende-se, em nosso estudo, às taxas de elocução. Nós observamos que os tempos entre respirações não se alteraram com as diferentes taxas de elocução, porque os falantes com taxas de elocução maiores inseriram mais material linguístico entre respirações consecutivas, quando comparados com falantes com taxas de elocução menores. Ou seja, nossos dados reforçam os dados da literatura no sentido de haver uma certa estabilidade, em cada falante, dos tempos entre respirações na fala lida.

Não se pode perder de vista, na comparação de nossos dados com os dados dos estudos citados, que estamos lidando — entre outros aspectos a se considerar — com línguas diferentes, de características prosódicas diferentes (BARBOSA, 2019), e com taxa de elocução (em sílabas/segundo) e densidade informacional diferentes (COUPÉ; OH; DEDIU; PELLEGRINO, 2019).

5.1.3.7 Incongruências segmentares vs. suprasegmentares

Todas as incongruências identificadas foram classificadas em segmentares ou suprasegmentares. Os fundamentos para a classificação de uma incongruência em segmentar ou suprasegmentar foram dois. O primeiro, de ordem teórica, baseou-se no conceito, já estabelecido na literatura, de que as línguas orais se compõem de elementos segmentares (no nível da palavra) e suprasegmentares (prosódia e outros fenômenos não segmentares); e não vimos razão para nos desviarmos dessa abordagem. O segundo fundamento foi de ordem prática. As incongruências

segmentares são bastante evidentes e objetivamente detectadas, pois dispomos de um "gabarito" — o texto escrito — para avaliar o desempenho dos falantes. Qualquer desvio desse gabarito indica, semiautomaticamente, uma incongruência segmentar. No caso das incongruências suprasegmentares, o cenário é diferente. Não dispomos de gabarito, pois os fenômenos prosódicos não estão codificados graficamente no texto escrito; são produzidos pelos falantes à medida em que vão lendo, segundo sua própria interpretação do que estão lendo, e de acordo com a impressão que querem causar em quem os ouve. Grosso modo, poderíamos dizer que, no primeiro caso (interpretação do que estão lendo), observamos a prosódia linguística e, no segundo caso, a prosódia afetiva voluntária. Em nosso estudo, restringimo-nos a tentar identificar as incongruências suprasegmentares relacionadas à prosódia linguística, porque, de acordo com nossas hipóteses, essas incongruências deveriam refletir equívocos dos falantes/leitores na interpretação da estrutura sintática e semântica dos textos escritos.

Parece-nos relevante acrescentar que a identificação e a classificação das incongruências suprasegmentares não foi uma tarefa simples. Como não temos um "gabarito", é inevitável que haja certo grau de subjetividade nesse processo. A dificuldade não está, propriamente, na identificação de uma manifestação prosódica incongruente. A simples percepção de que se trata de uma fala lida incongruente, e não de uma fala espontânea, é bastante imediata. Em nosso corpus, os oradores têm a intenção de ler como se estivessem falando espontaneamente, pois os textos escritos são apenas um suporte para uma manifestação política que se pretende oral. As incongruências, então, arruínam completamente essa intenção e são percebidas com muita facilidade. Tivemos, mesmo, a impressão — que não foi quantificada, nem qualificada e, portanto, permanece no campo subjetivo — de que as incongruências suprasegmentares tornam-se mais evidentes quando os falantes tentam soar mais naturais e espontâneos. Nossa hipótese não testada é que o uso de uma prosódia mais expressiva evidencia ainda mais os eventos de incongruência suprasegmentar.

A dificuldade não residiu, portanto, em apontar que houve desvios de prosódia, mas, sim, na classificação e interpretação desses desvios. Ou seja, percebe-se facilmente que determinada leitura não representa muito bem o que está escrito, mas, muitas vezes, não se sabe ao certo o que está errado. Uma solução definitiva para isso talvez resida em estudo com desenho experimental que controle todas as variáveis, mantendo invariáveis o falante e o material linguístico, e variando apenas o modo de emissão: leitura versus emissão oral. Não obstante, consideramos que, mesmo com tais limitações, nosso estudo foi capaz de identificar eventos incongruentes e correlacioná-los com determinadas variáveis.

Como vimos, foram identificados 2221 eventos de incongruência nas 5 horas de falas lidas. Quase metade desses eventos (45%) corresponderam a apenas três tipos de incongruências do tipo segmentar: trocas de palavras, desvios de ortoépia e desvios de concordância. Se incluirmos as próximas colocadas no ranking de ocorrências, que são incongruências suprasegmentares — pausa e inserção de fronteira não terminal —, teremos praticamente 70% de todas as incongruências encontradas. E, se acrescentarmos as três seguintes (acréscimo ou omissão de palavras, e inserção de fronteira terminal), teremos praticamente 90% das incongruências do corpus.

Em resumo, 90% das incongruências identificadas são dos tipos:

- Segmentares
 - troca/acréscimo/omissão de palavras
 - desvios de ortoépia
 - desvios de concordância
- Suprasegmentares
 - pausas
 - inserções de fronteira (não terminal ou terminal)

No cômputo geral, a maioria das incongruências (61%) foram incongruências segmentares. Isso não significa, entretanto, que elas sejam mais impactantes do que as incongruências suprasegmentares, pois o prejuízo à fala lida pode ser maior quando o falante insere ou exclui uma fronteira prosódica relevante do que quando troca uma palavra ou erra uma concordância. Como já discutimos, trata-se de categorias diferentes de incongruência. As incongruências segmentares são de caráter local, relacionam-se a palavras realizadas graficamente nos textos lidos e, portanto, estão imediatamente disponíveis para o leitor. Para evitar essas incongruências, basta ler as palavras da forma como elas foram escritas. São simples incongruências de leitura, apesar de sofrerem possível influência de outros fatores, como discutiremos a seguir. Além disso, como vimos, às vezes não causam prejuízo significativo e, quando causam, estão frequentemente associadas a retratações.

As incongruências suprasegmentares, por outro lado, formam uma categoria menos homogênea. As pausas, que são as mais frequentes nessa categoria, devem indicar uma dificuldade de processamento (que pode ser de ordem cognitiva, mas também pode ser, por exemplo, uma dificuldade puramente visual ou articulatória) relacionada a aspectos segmentares (palavras), suprasegmentares (prosódia), sintáticos, semânticos ou a quaisquer outras interferências que componham o universo

cognitivo do leitor naquele momento. As incongruências suprasegmentares de fronteiras, por outro lado, poderiam ser consideradas como uma categoria em si mesma, pois resultam de uma dificuldade de compreensão da estrutura do período, envolvendo conjuntos de palavras e a própria estrutura sintática e semântica do texto. São incongruências que advêm das próprias limitações da modalidade escrita, que, como já mencionamos, não é isomórfica com relação à modalidade falada (LYONS, 1982). Na sequência, discutiremos nossos achados com relação às incongruências segmentares e suprasegmentares, e exploraremos mais a fundo as ideias que expusemos acima.

5.2 As incongruências segmentares

Identificamos 1363 incongruências segmentares no corpus, sendo as mais frequentes, compondo três quartos dos casos, os desvios de ortoépia, as trocas de palavras e os desvios de concordância. Observamos que as incongruências segmentares poderiam ser subdivididas em dois grandes grupos, conforme a incongruência envolvesse a palavra inteira, ou operasse apenas no "interior" da palavra. Sob essa ótica, trocas de palavras estariam no primeiro grupo, e desvios de ortoépia e concordância, no segundo.

Essa subdivisão parece ter certa relevância, pois envolve aspectos diferentes do processamento de sentenças. Nós denominamos o primeiro grupo, arbitrariamente, como "incongruências frasais", pois operam no nível da frase, do conjunto de palavras, por meio de trocas, acréscimos, omissões ou repetições das unidades semânticas "autônomas" mínimas da língua, que seriam (não sem controvérsia) as palavras inteiras. As incongruências frasais, portanto, não "estragam" palavras, mas frases/períodos. Ao segundo grupo, demos o nome de "incongruências lexicais", pois operam no interior das unidades lexicais, as palavras, e incluímos nesse grupo as incongruências de ortoépia, concordância e repetição de sílabas (esta última, inclusive, poderia ser considerada um subtipo da classificação "ortoépia").

É interessante notar que essa subdivisão das incongruências segmentares se reparte aproximadamente ao meio, com metade das incongruências segmentares em cada grupo: lexical e frasal. Caberia, em nossa opinião, uma análise comparativa do comportamento de cada um desses grupos, para entender melhor se são fenômenos realmente distintos. Não o fizemos neste estudo porque só viemos a notar essa possível divisão conceitual do grupo das incongruências segmentares ao analisar os resultados obtidos, e não havia mais tempo hábil para tanto.

Quanto às incongruências de concordância, o achado de que o fenômeno predominante foi a singularização de palavras (que, pelas regras de concordância do PB, deveriam estar no plural) não indica que a singularização seja o fenômeno com a maior propensão para ocorrer dentre as diversas possibilidades de incongruência de concordância. Como não analisamos o corpus para determinar todas as situações em que poderiam ocorrer incongruências de concordância (todas as palavras no plural, singular, masculino ou feminino em situação gramatical de concordância com outra(s)), não sabemos se a singularização tem mais chance de ocorrer em cada situação particular de exigência de concordância. Por exemplo, em "os garotos pularam" temos 3 palavras e 7 "oportunidades" de errar a concordância, considerando as somas das combinações simples de um, dois e três "erros": 1) os garotos pulou; 2) os garoto pulou; 3) os garota pularam; e assim por diante. Pode ser que tenhamos "x" oportunidades de ocorrência de singularização e "x/10" oportunidades de ocorrência feminilização e, mesmo assim, tenhamos mais do que um décimo de feminilizações com relação a singularizações, o que indicaria que, em termos absolutos, tivemos mais singularizações, mas, em termos relativos, existiria uma chance maior de ocorrência de feminilizações em cada situação em que ela poderia ocorrer. Ou seja, haveria uma maior "pressão" para feminilizar do que para singularizar. Relembrando, não fizemos essa análise.

A singularização (nominal e verbal) como desvio da concordância normativa é um fenômeno bastante estudado em PB e em outras variedades do português. Brandão e Vieira (2012, p. 1061) observam que "no PB (...), o uso ou não da concordância canônica é uma opção que está claramente no nível de consciência dos indivíduos, de modo que é interpretado e funciona como forte traço caracterizador de classes sociais." E relata que a não marcação (não concordância) de plural é bastante variável geograficamente e socialmente, chegando a 22% de não marcação da concordância verbal em Nova Iguaçu (PB), por exemplo. E, de fato, é exatamente o que observamos em nosso cotidiano, onde "nós *tem* a impressão de que os *plural tá* morrendo *tudo*". E não vemos razão gramatical ou de processamento de linguagem para que não morram, pois são, muitas vezes, claramente redundantes. Mas devemos apontar que os desvios de concordância que observamos em nosso estudo não foram intencionais; não se tratou de uma opção consciente dos falantes; são incongruências involuntárias. E devem ter sido facilitadas por fatores linguísticos que, infelizmente, não analisamos, pois não foram incluídos em nossa proposta inicial de trabalho; mas podem, e mereceriam, ser analisadas no corpus que estudamos.

Outro fenômeno observado — que não foi antevisto na preparação teórica do estudo — foi a ocorrência de incongruências que não causaram prejuízo considerável ao período; ou seja, quem ouve

a fala lida sem ter tido acesso prévio ao texto escrito não é capaz de perceber que houve um desvio. Esse fenômeno ocorreu em 15% das incongruências segmentares, e sua distribuição foi desigual entre os diversos tipos de incongruência segmentar: foi de praticamente 50% nos acréscimos de palavras, a 0% nos desvios de ortoépia.

No caso do acréscimo de palavras, seria justo questionar até que ponto um acréscimo que não prejudica o período não seria totalmente intencional e, portanto, não poderia ser excluído da classificação de incongruências. Parece realmente mais provável que esses acréscimos tenham sido intencionais, mas não podemos afastar a hipótese de que possam ter sido involuntários e que o sucesso da empreitada tenha sido inteiramente fortuito. Nos casos em que houve prejuízo, o acréscimo também pode ter sido intencional, mas foi malsucedido e não há dúvidas de que deve ser considerado como incongruência. Trata-se de possibilidades que não investigamos e que poderiam ser objeto de estudos posteriores, até porque são fenômenos que também ocorrem nas outras categorias de incongruências segmentares frasais (omissão e troca de palavras), apesar de ocorrerem com frequências menores (23% e 17%, respectivamente).

5.2.1 As retratações

Restringimos nossas análises subsequentes aos casos em que houve prejuízo ao período — ou seja, 85% das incongruências segmentares —, justamente para evitar as possibilidades que discutimos acima. Se houve prejuízo, não há dúvida de que se trata de incongruência e de que sua produção foi, de alguma forma, involuntária. Notamos que, em muitos casos, os falantes preocupam-se em corrigir suas incongruências, retratando-se. E já mencionamos que as retratações ocorreram em 26% dos casos de incongruências em que houve prejuízo ao período, e que foram bem mais frequentes nos casos de desvio de ortoépia, em que metade das incongruências foi retratada. Mais adiante, nesta discussão, levantaremos e discutiremos a hipótese de que as retratações predominam neste tipo específico de incongruência segmentar porque o mecanismo de processamento da linguagem utiliza o resultado da produção da língua para compreender o que se pretendia dizer; ou seja, o produto linguístico só pode ser monitorado após ser produzido.

O fenômeno de retratação traz consigo outro fenômeno interessante, que é a retratação com adição de palavras. Em 36% dos casos de trocas de palavras e desvios de ortoépia (as incongruências segmentares mais comuns), o falante acrescenta à retratação da palavra incongruente uma ou mais palavras anteriores cuja emissão não foi incongruente.

Os dados foram descritos no capítulo Resultados, e o que se deduz é que a adição de palavra congruente à retratação de uma palavra incongruente (ex.: “caí no buraco...no buraco”) ocorre:

- com várias categorias gramaticais
- com adição de até 3 palavras congruentes à retratação da palavra incongruente
- em situações onde a palavra incongruente não é muito grande (até 4 sílabas) e a palavra adicionada é pequena (1 sílaba)
- mas, sobretudo, com forte predomínio, em ordem decrescente de incidência, das combinações:
 - artigo definido + substantivo
 - contração preposição/artigo definido + substantivo
 - preposição + substantivo

Os substantivos incongruentes são retratados, na maioria das vezes, acompanhados pela palavra imediatamente anterior, que costuma ser artigo definido, ou preposição+artigo definido, ou preposição. Por exemplo, é muito comum um padrão de retratação como em "a janel...a panela..". E nos perguntamos: não bastaria dizer "a janel...panela..."? Os padrões observados em nosso corpus sugerem que, por alguma razão, não basta. Ainda não temos clareza quanto às motivações do fenômeno, mas vamos atacá-lo por dois flancos, seguindo duas abordagens teóricas diferentes e complementares: a noção de "clíticos"; e a discussão sobre a universalidade do sintagma determinante (*determiner phrase*, DP) em prejuízo do sintagma nominal (*nominal phrase*, NP).

De acordo com Basílio (2004, p. 14), clíticos são "unidades que se agregam a uma palavra fonologicamente, sem fazer parte dela do ponto de vista morfológico", e estariam nessa categoria, em PB, além dos pronomes pessoais átonos, a classe dos artigos. A autora também diz que a existência dos clíticos torna mais difícil a identificação dos limites da palavra, pois a associação clítico + palavra compõe uma unidade fonológica (palavra fonológica), apesar de não compor uma unidade (palavra) morfológica. A combinação artigo definido + substantivo, que mencionamos no parágrafo anterior, corresponde à combinação de um clítico com uma palavra lexical "autônoma". O artigo definido, portanto, estaria fonologicamente "apoiado" no substantivo e os dois comporiam uma palavra fonológica, como em "a panela". De acordo com Zwicky (1985), haveria uma gradação (ou um contínuo) das unidades das línguas, seguindo a ordem afixo-> clítico-> palavra-> frase-> oração. Afixos e clíticos, portanto, seriam partículas (o autor não aconselha o uso desse termo, que considera enganador) sem autonomia completa.

Segundo esses conceitos, o artigo definido que é adicionado à retratação do substantivo não teria autonomia e não poderia existir sem o apoio fonológico do substantivo. Talvez por isso, pelo conjunto artigo definido + substantivo compor uma palavra fonológica, uma unidade prosódica, o falante faz a retratação da unidade fonológica completa, e não apenas do item lexical, da palavra gráfica incongruente (substantivo). No entanto, tivemos poucas oportunidades de observar como seria a retratação em caso de incongruências de artigos definidos (apenas 3 incongruências dos tipos troca de palavras e ortoépia: uma com adição de palavra, e duas sem) e, mesmo que tivéssemos mais oportunidades, o fato de ele anteceder o substantivo em PB não favoreceria nossa análise de que ele demandaria o substantivo na retratação. Observamos, entretanto, que os artigos indefinidos, ao contrário dos definidos, não participaram do fenômeno de retratação com adição de palavras, sendo adicionados ou não adicionados à retratação com igual frequência. Pode-se argumentar que a dependência fonológica de artigos indefinidos em PB seria menor, que eles teriam mais autonomia, mas, se estamos em busca de um padrão, essa dissonância no comportamento da retratação entre tipos de artigos não favorece muito a hipótese de que o conceito genérico de “clíticos” estaria relacionado à retratação.

É bem provável que, na mente do falante/leitor, a combinação artigo definido + substantivo seja considerada como uma unidade, e não como uma sequência de duas palavras autônomas. O falante produz uma incongruência segmentar na emissão do substantivo, percebe que produziu uma incongruência, e a corrige. Mas, em vez de retificar apenas o substantivo, faz uma retificação que inclui o artigo definido, como se a incongruência tivesse sido produzida na unidade artigo definido + substantivo, e não na unidade lexical substantivo, apenas. Sob essa ótica, "a janela" não seria "a"+"janela", mas "ajanela", um "sintagma determinante", em oposição a "janela", um "sintagma nominal". O contraste pode ser apreciado em "a janela custou caro" vs. "janela custa caro".

A interpretação do padrão de retratação que estamos discutindo parece tangenciar o debate sobre as hipóteses do *Universal-DP* vs. *Parametrized-DP*, ou sintagma determinante universal vs. sintagma determinante parametrizado (CYRINO; ESPINAL, 2014). Os propositores da primeira hipótese defendem a universalidade do sintagma determinante, mesmo em línguas que não têm artigos expressos; e os propositores da segunda hipótese defendem que línguas sem artigos (como o russo) parametrizam os constituintes nominais como sintagmas nominais, apenas (sem a necessidade de imaginar sintagmas determinantes profundos com determinantes superficiais nulos, não expressos).

Como se sabe, o PB e as diversas línguas latinas são exemplos interessantes do surgimento diacrônico de determinantes (como artigos, por exemplo) a partir de uma língua que não os continha (o latim) (SILVA, 2020). O PB, portanto, é uma língua que possui artigos e que, de acordo com aqueles que militam no campo do gerativismo — e que buscam uma coerência sintática universal que justifique a defesa de um módulo sintático cerebral genético único em todo ser humano —, projeta um sintagma determinante (DP). A adoção do conceito de um sintagma determinante implica que o núcleo do sintagma "a janela" seja o artigo definido. Sob esse ponto de vista, poderíamos justificar o fenômeno de retratação com adição de palavra em situações do tipo artigo definido + substantivo como um reflexo dessa organização sintática profunda de nossa língua, em que o substantivo, quando determinado, seria um item lexical secundário (não nuclear) no sintagma determinante. Nessa perspectiva, a retratação do substantivo exigiria a inclusão do artigo definido, pois o objetivo seria retratar, na verdade, o constituinte, o DP.

No entanto, acreditamos ser possível (ou gostaríamos que fosse) incorporar o fenômeno da retratação com adição de palavras ao rol dos fatos linguísticos pendentes de explicação sem, necessariamente, explicá-lo por meio de conceitos gerativistas. O fato de que um substantivo determinado por um artigo constitua uma unidade semântica diferente de um substantivo não determinado não implica a existência de um DP universal, e nem mesmo de um DP. E não implica, tampouco, a existência de uma sintaxe única derivada de um módulo gramático universal inscrito nos cérebros humanos. Significa apenas que, ao dizermos "a panela", estamos querendo dizer "a panela", e não "panela". Significa que não estamos satisfeitos em dizer "a janel...panela...", mas que consideramos ser necessário (inconscientemente, o que é mais provável) retratar a unidade que representa a ideia que queremos transmitir, que é a de um substantivo determinado por um artigo definido ("a panela"); e que essa unidade, essa ideia, não pode ser representada apenas pela palavra "panela". E, muito provavelmente, porque, em PB, se dizemos "panela", estamos dizendo algo bem diferente de "a panela". Essa discussão, em nossa opinião, mereceria um aprofundamento que, por razões já expostas, não faremos.

5.2.2 As trocas de palavras

5.2.2.1 A frequência das palavras nos corpora

Nossa principal hipótese com relação às trocas de palavras era a de que as palavras inseridas — que substituem, incongruamente, as palavras originais escritas no texto — seriam mais

frequentes na língua do que as palavras originais. Com base em autores como Langacker (2017) e Pickering e Garrod (2013), acreditávamos que palavras mais usadas pelos falantes do PB (mais frequentes) estariam cognitivamente mais disponíveis e, portanto, teriam mais chance de serem escolhidas em substituição incongruente às palavras originais, presentes nos textos escritos.

Nosso parâmetro para atribuir frequências de ocorrência às palavras do nosso corpus foi um corpus de PB — CB (SARDINHA; MOREIRA FILHO; ALAMBERT, 2019) — composto por cerca de 1 bilhão de palavras. Uma análise das 19 palavras mais frequentes em nosso corpus (**Tabela 31** – Número de trocas vs. Ocorrências nos textos vs. Ocorrências no CB) mostrou que as palavras mais frequentes dentre as 33860 analisadas em nosso corpus são praticamente as mesmas que se mostraram mais frequentes no corpus CB. Esse fato, por si só, já é bem interessante, pois nosso corpus, além de ser relativamente pequeno, tem características bastante específicas, como sua finalidade (o convencimento político), sua linguagem mais formal, e seus temas (que, apesar de variados, são limitados ao espectro de atuação do parlamento).

Essa concordância entre nosso corpus e o CB quanto às palavras mais frequentes, ao que parece, deve-se ao fato de que as palavras mais frequentes em PB (e, arriscamos dizer, em qualquer língua humana oral) sejam palavras gramaticais, como preposições, artigos, pronomes e conjunções. Somadas a essas — e, neste caso, não arriscamos falar de outras línguas, mas apenas de PB —, encontram-se dois advérbios (*não*, *mais*) e o nosso verbo padrão para cópulas (*é*). Aparentemente, não se constrói uma sentença sem utilizar algumas dessas palavras e, dessa forma, elas se fazem presentes em praticamente todos os tipos de construções linguísticas em PB. O contraste entre essas palavras e a categoria de palavras que Payne (2006) denomina "palavras lexicais plenas" (*full lexical words*) é evidente. As "palavras lexicais plenas" são classes abertas, continuamente atualizadas e, portanto, incluem um número muito maior de itens lexicais; mas cada um desses itens, em particular, tem uso individual bem menos frequente do que o de palavras gramaticais.

Nossas análises das frequências das palavras originais (escritas no texto) e inseridas (faladas) trouxeram resultados bem menos diretos do que esperávamos. Uma variável insuspeita — a retratação — teve de ser introduzida na equação. De fato, as palavras inseridas são mais frequentes (na língua) do que as palavras originais, como previmos, mas essa diferença só é observada nos casos em que a troca de palavras é seguida por retratação.

Por exemplo, a frase escrita “...o conjunto de mulheres e homens em certa pátria...” foi lida como “...o conjunto de mulheres e homens **de...em** certa pátria...”. Houve troca de “em” por “de”, com retratação posterior. Neste caso, “de” é quatro vezes mais frequente na língua do que “em” (“de”, na realidade, é a palavra mais frequente do PB, segundo o CB). Na frase escrita “...em última instância, razão e sensibilidade podem...”, lida como “...em última instância, razão **da** sensibilidade podem...”, trocou-se “e” (...“razão e sensibilidade”...) por “da”, sem retratação, sendo que “da” é duas vezes menos frequente na língua do que “e”. O contraste entre essas duas frases ilustra o cenário que observamos em nossa análise: trocas por palavras mais frequentes que as palavras originais associaram-se a retratação; trocas por palavras com frequência semelhante não se associaram a retratação.

Analisando melhor esse fenômeno, observamos que a diferença entre as frequências das palavras originais e inseridas nos casos com retratação deveu-se à frequência das palavras originais, e não das inseridas. O que diferencia as trocas com retratação das trocas sem retratação é que as palavras originais, no caso das trocas com retratação, são menos frequentes na língua, sugerindo que a variável que se associa à retratação é a frequência absoluta da palavra original, e não a relação entre as frequências de palavras originais e inseridas.

Como não encontramos na literatura descrição ou análise desse fenômeno, levantamos algumas hipóteses e escolhemos a menos problemática. Tal hipótese exigiria uma reanálise dos dados obtidos que, por exiguidade de prazo, não foi realizada. A ideia é que os mecanismos de processamento de linguagem envolvidos nas trocas com retratação seriam (quantitativamente) diferentes daqueles envolvidos nas trocas sem retratação. No primeiro caso, teríamos uma influência maior da baixa frequência relativa da palavra original como mecanismo de troca; no segundo, teríamos uma influência maior do contexto semântico/gramatical do texto.

Sob essa perspectiva, o fator responsável pela ocorrência da retratação seria o mecanismo de troca, e não a frequência das palavras. Trocas induzidas pelo contexto seriam menos “aberrantes”, menos perceptíveis ao falante, e, portanto, passariam despercebidas e não seriam retratadas. Trocas induzidas pelo grau de “surpresa” da palavra original — uma palavra que, por ser menos frequente, estaria cognitivamente menos disponível — seriam mais evidentes, pelo fato de serem erros pontuais, locais, menos influenciados pelo contexto; e, por isso, mais facilmente percebidas e retratadas.

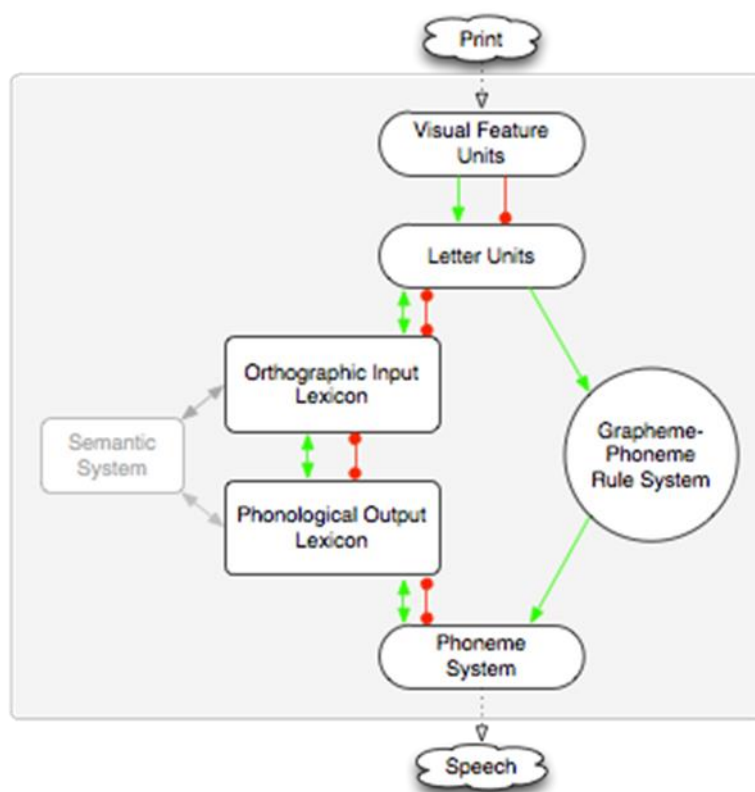
5.2.2.2 As semelhanças entre palavras originais e inseridas

Apesar de não termos confirmado de forma incondicional nossa hipótese inicial de que a frequência da palavra na língua seria um fator onipresente nas trocas de palavras — já que ela se mostrou relevante apenas nos 32% dos casos em que a troca se acompanhou de retratação —, não poderíamos encerrar a discussão sem voltar a mencionar teorias que tratam do *predictive language processing*, como as defendidas, entre outros, por Pickering e Garrod (2013). Antes de tratar disso, entretanto, precisamos discutir outro fenômeno que observamos nas trocas de palavras: a semelhança entre palavra original e palavra inserida. Observamos que na grande maioria dos casos (78%) a troca é feita por uma palavra que tem semelhança gráfica/fonológica com a palavra original. O padrão se aplica, principalmente, às palavras lexicais, em que 92% das trocas são feitas por palavras semelhantes. Além disso, em metade dos casos (predominando entre as palavras gramaticais), a semelhança inclui o compartilhamento de lexema.

O que se percebe é que uma palavra, via de regra, é trocada por um parônimo, uma palavra parecida e, muitas vezes, essa semelhança envolve um compartilhamento de lexema. O que resta saber é qual fator seria mais importante nesse fenômeno: a) o fato de haver um lexema comum, uma associação semântica entre a palavra original e a palavra inserida; ou b) o simples fato de haver uma semelhança fonológica. Palavras que compartilham lexemas, além de serem semanticamente relacionadas, também são parônimos, palavras parecidas, pois, via de regra, compartilham a mesma raiz. Sendo assim, não conseguimos segregar o grau de influência de cada componente — semelhança fonológica vs. semelhança semântica — sobre a escolha da palavra inserida. É possível que a semelhança semântica não tenha nenhuma influência sobre a troca e, nesse caso, a troca seria feita apenas com base nos aspectos fonológicos da palavra. Análises adicionais que poderiam ter sido feitas seriam: a) classificar os pares de parônimos em homógrafos e/ou homófonos, o que poderia indicar alguma dissociação entre características fonológicas e características puramente gráficas; e b) analisar possíveis associações semânticas entre as palavras originais e inseridas seguindo protocolos como os propostos por Calais, Lima-Gregio, Arantes, Gil e Borges (2016), que identificaram diferentes graus de associação semântica entre as palavras, justificando, talvez, parte das “escolhas” das palavras inseridas por seu grau de associação com a palavra original. Entretanto, apesar de não termos analisado associações semânticas, há indícios de que, no processo de leitura, essas associações poderiam ter um caráter secundário.

As possibilidades discutidas acima são bastante significativas para a compreensão dos mecanismos de processamento da linguagem e encontram amparo em um modelo de leitura em voz alta que defende uma dupla rota de processamento para a leitura (*dual-route approach*), descrito em Coltheart (2005) e ilustrado abaixo.

Figura 31 - Modelo *dual-route approach* de processamento de leitura



A **Figura 31** (obtida em: <https://maxcoltheart.wordpress.com/drc/>) sugere que as palavras gráficas não precisam alcançar o sistema semântico para ativarem o léxico e o sistema fonológico; ou seja, poderiam dar origem à articulação sonora da palavra gráfica sem, necessariamente, ativar seu significado. É um modelo compatível com o que mencionamos no capítulo Revisão da Literatura, onde dissemos que, segundo Dehaene (2009, figura 2.18 no texto original), "após o processamento inicial em áreas visuais, há uma bifurcação condicionada pela tarefa imposta ao leitor: se a tarefa privilegia a compreensão do texto, ou, em termos mais específicos, o acesso ao significado das palavras, a rota passa pelas regiões inferiores do lobo temporal; se a tarefa privilegia a produção correta dos sons da fala, a rota segue por regiões mais superiores do mesmo lobo". Em um modelo como esse, é

perfeitamente factível que, em alguma medida, a semelhança gráfica entre duas palavras seja a única responsável pela substituição de uma palavra (a original) por outra (a inserida).

Outro modelo que também considera uma certa independência cognitiva entre o sistema semântico e o fonológico foi proposto por Levelt (1983, p. 2), que argumentou que "*speakers have little or no access to their speech production process; self-monitoring is probably based on parsing one's own inner or overt speech*"¹. Com base em análises de erros e retratações de fala, o autor conclui que existem duas vias de monitoramento da fala: uma interna, que monitora a fala interior, silenciosa (*inner speech*); outra externa, que monitora a fala exterior, sonora. E, adicionalmente, que o monitoramento da produção incidiria sobre os produtos — as palavras —, e não sobre os processos — os mecanismos semânticos e sintáticos.

Pickering e Garrod (2013) observam que a maioria das abordagens de processamento de linguagem considera a existência de dois módulos independentes, um dedicado à produção, outro à compreensão; mas que o modelo de Levelt (1983), ao prever mecanismos de monitoramento da produção da fala, propõe que compreensão e produção são fenômenos integrados. Eles apoiam essa integração e propõem um modelo em que há monitoramento da produção por mecanismos de compreensão, e vice-versa. Em contraste com a proposta de Levelt (1983), eles propõem três modificações: haveria um mecanismo de produção antecipada de linguagem, em paralelo com o mecanismo de produção usual; existiria um módulo dedicado exclusivamente ao monitoramento, comparando o resultado das produções linguísticas entre os dois mecanismos de produção (o usual, e o de produção antecipada); e o monitoramento aconteceria tanto em nível fonológico, quanto sintático e semântico. Este modelo de Pickering e Garrod (2013) destaca-se por: a) defender a existência de um mecanismo de antecipação da produção da linguagem integrando produção e compreensão; b) alinhar-se aos modelos, como o de Levelt (1983), que acreditam em uma via interna (*internal loop*) de monitoramento da produção; e c) incluir um monitoramento semântico anterior ao monitoramento fonológico.

Tais ideias são criticadas por Hartsuiker (2013), que não acredita na necessidade de um módulo dedicado à produção antecipada da linguagem. Segundo ele, tal desenho implicaria desperdício de

¹ Os falantes têm pouco ou nenhum acesso aos processos de produção da fala; o automonitoramento deve se basear na análise da própria fala interna (silenciosa) ou externa (em voz alta). (tradução nossa)

recursos cognitivos, já que os mecanismos tradicionais de produção já seriam suficientes para executar a tarefa. Ele critica, também, a ideia de um monitoramento interno, anterior à emissão sonora, e descarta a possibilidade de um monitoramento semântico anterior à produção da fala. Em suma, ele descarta quase que completamente as propostas anteriores. Essas críticas baseiam-se em dois estudos que se utilizam do paradigma de "rastreamento ocular em mundo visual" (*visual world eye-tracking*). Nesse paradigma experimental, o participante olha para uma tela com algumas imagens e a fixação do seu olhar em cada uma das imagens é continuamente monitorada. No primeiro estudo (HUETTIG; MCQUEEN, 2007), os participantes ouviam determinadas palavras enquanto olhavam para uma tela onde outras palavras estavam escritas. Observou-se que as fixações eram mais frequentes em palavras fonologicamente semelhantes à palavra ouvida. Por exemplo, eles ouviam a palavra holandesa *beker* (béquer) e olhavam para *bever* (castor), em vez de olhar para *klos* (carretel), *vork* (garfo) ou *paraplu* (guarda-chuva). No segundo estudo (HUETTIG; HARTSUIKER, 2010), em vez de 4 palavras na tela, havia 3 palavras e uma imagem. O participante tinha de nomear a imagem em voz alta. Ao nomear a imagem de um coração (*hart*), fixava mais o olhar na palavra gráfica *harp* (harpa), fonologicamente semelhante, do que em *zetel* (sofá) ou *raam* (janela).

Analisando o tempo entre o início da emissão sonora da palavra (ouvida, no primeiro estudo, e pronunciada a partir da nomeação de uma imagem, no segundo) e a fixação do olhar na palavra gráfica fonologicamente semelhante, os estudos do parágrafo anterior observaram que eles eram idênticos nas duas condições experimentais. Os autores argumentaram que, se houvesse uma via interna (*internal loop*) de monitoramento da produção da linguagem, o tempo para fixação do olhar na palavra fonologicamente correlata deveria ser menor na segunda condição experimental, em que o participante olha para a imagem e produz o seu nome. Nesse processo, segundo os modelos que defendem uma via interna de monitoramento fonológico, deveria haver monitoramento (percepção) da fala interior, induzindo uma fixação mais precoce do olhar nas palavras fonologicamente correlatas, o que não ocorreu.

Podemos acreditar, portanto, que as retratações de incongruências segmentares são o resultado de um processo em que o falante deve seguir, grosso modo, alguns passos: ver a palavra -> produzir seu correspondente fonológico em voz alta -> detectar uma incongruência ao ouvir a própria fala -> retratar-se. As retratações ocorreriam quando o falante percebesse a incongruência ao ouvir a própria voz. E, seguindo esse raciocínio, seria possível levantar hipóteses para explicar a maior ou menor incidência de retratações entre as diversas categorias de incongruências segmentares observadas em nosso estudo. As hipóteses, agrupadas pela incidência de retratação, seriam as seguintes:

- Menos retratadas:
 - acréscimos de palavras: apesar de incongruentes, eles devem ter sido, em muitos casos, voluntários e, portanto, o falante não viu necessidade de retratação porque a palavra teria sido inserida de propósito
 - omissão de palavras: considerando que é necessário ouvir a palavra falada para retratar-se, pode-se acreditar que, ao não ouvir o erro, o falante não se retrata
 - desvios de concordância: a hipótese mais provável relaciona-se ao contraste entre a norma padrão e os usos correntes da língua, associados a suas variações regionais. Os desvios caracterizados por pluralização, masculinização e feminilização são incongruências evidentes e involuntárias, mas, no total, correspondem a pouco menos de metade da categoria predominante, que é a singularização. E a singularização (em desacordo com a norma padrão) é de uso cada vez mais comum na língua do dia a dia. Assim, é bem provável que os falantes que singularizaram palavras durante a leitura também singularizem palavras em situações informais; portanto, ao ouvirem as palavras singularizadas, não notaram a dissonância com o texto escrito. Ou, até mesmo (o que considero bem pouco provável), singularizaram voluntariamente
- Mais retratada:
 - desvio de ortoépia: deve ser mais retratado porque chama mais atenção, é uma incongruência mais evidente aos ouvidos do falante; incomoda mais, e não deixa muita margem a uma negligência voluntária
- Prevalência intermediária de retratação:
 - troca de palavras: essa categoria está numa situação intermediária; podem ser notadas, ou não; se notadas, o falante pode optar pela retratação, ou não; e, se não notadas, não serão retratadas; então, os 32% de retratações indicam trocas que foram notadas após serem articuladas verbalmente e em que o falante achou por bem retratar-se; e os 68% não retratados incluem as trocas não notadas e as notadas que não incomodaram o falante a ponto de serem retratadas

Observando esses resultados retrospectivamente, reconhecemos que caberia uma análise mais detida, comparando os grupos com e sem retratação dentro das categorias ortoépia e troca de palavras, o que poderia revelar outros aspectos associados à retratação que não ficaram evidentes na análise que fizemos.

5.2.2.3 Observações adicionais sobre as trocas de palavras

5.2.2.3.1 “Resistência” à substituição

Outro resultado que obtivemos foi que determinadas palavras sofrem trocas com frequência diferente de suas frequências de ocorrência nos corpora. Certas palavras parecem ser mais resistentes aos mecanismos que induzem as trocas, enquanto outras parecem ser mais suscetíveis. No primeiro grupo estão "que", "um", "uma" e "mais". No segundo, "por", "ao", "nos" e "das". Poderíamos levantar hipóteses para justificar esse achado, mas elas seriam bastante precárias. Se tivéssemos um corpus maior, com uma amostra maior de trocas de palavras, talvez observássemos outros casos de resistência ou suscetibilidade às trocas e pudéssemos, então, elaborar melhor essas hipóteses. Uma delas seria a de que palavras mais suscetíveis pertenceriam a categorias mais produtivas, com mais opções de palavras substitutas com função gramatical equivalente, como seriam as preposições presentes em "por", "ao", "nos" e "das". E, da mesma forma, palavras mais resistentes pertenceriam a categorias menos produtivas, com menos ou nenhuma opção de troca, como seriam "que", "um/uma" e "mais".

5.2.2.3.2 O que é que o que é?

Chamou-nos atenção o caso específico da palavra "que", que apareceu 895 vezes nos textos lidos e não sofreu uma troca sequer. Como foi um achado imprevisto, retornamos ao corpus e fizemos uma análise sintática/gramatical das formas de apresentação do "que". Escolhemos aleatoriamente 50 ocorrências do "que" e as classificamos gramaticalmente, mas não sem antes enfrentar um problema, ou, melhor dizendo, um incômodo. De acordo com livros-texto e gramáticas tradicionais, a palavra "que" desempenha múltiplas funções, que são, inclusive, conhecidas como "as funções do que". Segundo Sarmiento (2012, p. 464), o "que", dependendo do contexto, "pode ser advérbio, substantivo, preposição, interjeição, conjunção, pronome interrogativo adjetivo, pronome interrogativo substantivo, pronome relativo, pronome indefinido ou partícula expletiva". Cunha e Cintra (2001), por sua vez, o classificam, pelo menos, como pronome relativo, pronome interrogativo, advérbio comparativo, conjunção coordenativa e conjunção subordinativa. O dicionário Aurélio acrescenta às classificações acima as de pronome exclamativo e vários subtipos de conjunção coordenativa (alternativa, subordinativa comparativa) e subordinativa (integrante, concessiva, causal e final). E o dicionário Houaiss adiciona mais duas categorias: conjunção explicativa e conformativa. Será que existe, em PB, outra palavra que tenha tantas funções quanto o "que"?

Causa-nos estranheza que um item lexical tenha tantos e tão variados empregos na língua e que, a despeito disso, consiga exercer todas as suas funções sem servir como um fator de confusão para seus usuários. O fato de o "que", com 895 ocorrências, não ter sofrido nenhuma troca indica que, aparentemente, ele não tem substitutos à altura ou, pelo menos, não tem substitutos que sejam suficientemente semelhantes (fonologicamente ou graficamente) para serem confundidos com ele.

No CB, o "que" é a quarta palavra mais frequente. Só é superada pela preposição "de", pelos artigos definidos no singular e suas contrações com "de" (o, a, do, da), e pela conjunção "e". Ou seja, é uma palavra de uso extremamente frequente nas estruturas linguísticas do PB. Na literatura, as funções do "que" (e de seus correlatos em outras línguas, como o *that*, em inglês) que têm merecido mais atenção são as de relativizar (*relativizer*, que corresponderia ao nosso pronome relativo) e complementar (*complementizer*, que corresponderia à nossa conjunção integrante). Como é de praxe em questões linguísticas, há controvérsias, e uma delas foi proposta por Kayne (2010). Ele defende que o *that* do inglês tem apenas a função de relativizador e que as orações introduzidas por ele são sempre orações relativas (subordinadas adjetivas, em PB). Esse conceito, segundo o autor, estender-se-ia, com o perdão da mesóclise, a outras línguas, o que equivale a dizer que o "que" em PB teria, segundo Kayne (2010), apenas a função de pronome relativo, e não de conjunção integrante.

Essa ideia nos agrada por seu caráter unificador. Precisamos classificar 50 ocorrências do "que" em nosso corpus e não nos conformamos com o número de valores que, segundo a gramática normativa, essa variável poderia assumir. Somando o inconformismo a certas observações, e adicionando alguns dados objetivos, elaboramos algumas hipóteses. Nossa impressão é que a palavra multifuncional "que" não tem múltiplas funções, e talvez tenha apenas uma. E, tomando a liberdade de discordar de Kayne (2010), essa função não seria a de pronome relativo. Nessa perspectiva, funções adicionais eventualmente atribuídas ao "que" resultariam de locuções (da associação do "que" com outras palavras), ou de idiossincrasias idiomáticas, de provável justificativa diacrônica.

As 50 ocorrências aleatórias do "que" em nosso corpus apresentam, via de regra, o mesmo padrão estrutural: ele separa verbos. Ou seja, o "que" funciona como um marcador, um delimitador entre duas estruturas predicativas verbais. Dentre as 50 ocorrências, observamos apenas 4 exceções, que talvez não sejam exatamente exceções, e que são descritas abaixo:

- 2 casos com apenas 1 estrutura verbal intermediada e "regida" pelo "que":
 - "Época em *que* as mulheres **havam acabado** de conquistar o direito ao voto."

- Pode ser reescrita como "Aquela **era** uma + época em *que* as mulheres **havam acabado...**"
 - "*Que* o espírito do Rio **volte** a animar esta casa."
 - Pode ser reescrita como "**Quero** *que* o espírito do Rio **volte...**"
- 2 casos em que a estrutura regida pelo "que" não tem verbo:
 - O Fórum Mundial acaba realizando *mais que* um ranking.
 - Trata-se de uma locução, "mais que", e não de "que"
 - Isso é o *mesmo que* nada!
 - Trata-se de uma locução, "o mesmo que", e não de "que".

Em todos os demais casos, o "que" delimitou duas estruturas construídas ao redor de verbos. Em 39 ocorrências (excluídos os 2 casos que interpretamos como locuções), tínhamos a estrutura ORAÇÃO+QUE+ORAÇÃO; e em 9 ocorrências, tínhamos a estrutura QUE+ORAÇÃO+ORAÇÃO.

Essa observação, aparentemente, não acrescenta nada de novo ao que já é de conhecimento geral, pois é amplamente reconhecido que o "que" tem, como funções principais, a de relacionar/relativizar (pronomes relativos) e a de integrar (conjunção integrante) estruturas verbais, orações. No período abaixo, por exemplo, a primeira ocorrência do "que" seria classificada como conjunção, e a segunda, como pronome:

"Avisa lá **que** eu vou chegar mais tarde; vou me juntar ao Olodum **que** é da alegria."

Nossa suspeita, entretanto, é que o "que" não exerce nem a função de pronome, nem a função de conjunção. Não exerce a função de pronome porque a existência de um pronome numa situação como a da segunda ocorrência do "que" no período acima é pragmaticamente dispensável; já há, no período, as informações linguísticas necessárias e suficientes para que não seja necessário recorrer a teorias que propõem a existência de uma estrutura sintática completa subjacente que precisaria ser licenciada. E o "que" também não exerceria a função de conjunção porque não há necessidade de juntar ou integrar duas orações adjacentes, sequenciais, que já estão juntas e integradas pela própria serialidade e pelas relações de dependência da língua: "avisa lá: eu vou chegar mais tarde". Mas, se o "que", hipoteticamente, não integra nada, nem atua como substituto ou referência de nenhum nome, para que o "que" existe?

Nossa impressão — que não foi exaustivamente investigada neste trabalho, pois foge de seu escopo — é que o "que" funciona apenas e simplesmente como um delimitador, sem conteúdo

semântico próprio ou função adicional. Ele simplesmente delimita as áreas de atuação de duas estruturas linguísticas organizadas ao redor de verbos. Introduce-se o "que" em determinado momento do fluxo linguístico apenas para evitar que constituintes obrigatórios ou opcionais associados ao verbo A sejam confundidos com constituintes obrigatórios ou opcionais associados ao verbo B e, assim, facilitar o processamento da língua. Ele seria uma palavra gramatical simples, sem conteúdo semântico nenhum, sem nenhuma outra função, sem substitutos de uso frequente ou de semelhança fonológica/gráfica aparente e, por todas essas razões, não sofre (não sofreu) trocas por outras palavras durante 5 horas de leitura em voz alta. As diversas línguas, com o objetivo principal de indicar a fronteira entre os territórios de 2 verbos, introduziriam nessa fronteira um marcador, que poderia ser uma palavra (como o "que", em PB), ou algum outro mecanismo delimitador qualquer, como ordem de constituintes, eventos prosódicos etc. As demais funções tradicionalmente atribuídas, em PB, ao "que" seriam exercidas, na realidade, por quaisquer outras entidades linguísticas presentes no período, mas não pelo "que".

Há, em PB, fenômenos linguísticos que podem ser interpretados como sugestivos de que essa hipótese, eventualmente, poderia fazer sentido. Camacho (2013), em uma abordagem discursivo-funcional das construções relativas em português, lembra-nos dos estudos pioneiros de Fernando Tarallo e de estudos subsequentes diversos, como os de Mary Kato e outros, apontando e interpretando variantes não padrão de uso do "que". E fornece exemplos dessas variantes não padrão em orações relativas, chamadas, em contraste com a construção padrão (a), de copiadora (b) e cortadora (c):

- a) A menina de que você gosta mora na casa da esquina
- b) A menina que você gosta dela mora na casa da esquina
- c) A menina que você gosta mora na casa da esquina

Em (b), o emprego do "dela" revela o esvaziamento da referencialidade do "que", em consonância com a ideia de que ele seja apenas um delimitador, como argumentamos. E em (c), a ausência da preposição "de" — que não deixa de ser usada, no uso informal do dia a dia, quando o verbo "gostar" está em relação direta com seu objeto (ninguém diz, mesmo informalmente, "gostar ela", mas sim "gostar dela") — em associação com o "que", também revela o mesmo esvaziamento de sua função pronominal. Essas variantes não padrão revelam bastante sobre a real identidade das palavras da língua, em um uso não patrulado, mais livre e espontâneo. Reconhecemos a superficialidade e a precariedade na introdução dessa hipótese, mas, considerando os dados que obtivemos, não poderíamos deixar de mencioná-la, mesmo que apenas de passagem.

5.2.2.3.3 Tendência para ser escolhida como palavra substituta

A discussão sobre o "que" surgiu da observação de que ele foi a palavra original com a menor razão entre troca e ocorrência de palavra, pois foi utilizada 895 vezes e não foi trocada uma vez sequer. A partir dessa observação, procuramos descobrir se há, do outro lado, do lado das palavras inseridas, alguma que tenha maior tendência para ser escolhida como substituta da palavra original. Observamos que o "de" sofreu quase o dobro de inserções com relação ao número de substituições (foi mais frequente entre as palavras inseridas do que entre as palavras originais trocadas) e o "a" sofreu menos de metade de inserções do que de substituições (foi mais trocado do que inserido). Os testes estatísticos aplicados ao conjunto das palavras mais trocadas, entretanto, não considerou essas diferenças observadas como significativas. Apesar disso, parece plausível que o "de", como a palavra mais frequente do PB, esteja cognitivamente mais disponível para inserções do que palavras menos frequentes, o que poderia ser comprovado por meio de um estudo dedicado a essa questão específica e com número de ocorrências suficiente para superar os limites dos testes estatísticos.

5.2.2.3.4 O papel das categorias gramaticais nas trocas de palavras

Também analisamos a proporção das categorias gramaticais entre as palavras originais e inseridas. Classificamos as palavras em gramaticais e lexicais, e estas últimas em substantivo, adjetivo, advérbio, verbo ou particípio. O que observamos foi que houve troca de categoria gramatical em 17% dos casos de troca de palavras, sendo que a troca mais comum foi entre substantivos e adjetivos (e vice-versa). Se considerarmos o percentual de troca dentro de cada categoria gramatical, percebemos que as categorias mais sujeitas a troca (quando a palavra inserida é de categoria diferente da palavra original) foram advérbios, adjetivos e particípios.

O principal achado dessa análise, entretanto, foi que, a despeito de um percentual razoável de trocas de categoria (17%), a distribuição de categorias no conjunto das palavras inseridas permaneceu igual à das palavras originais. Ou seja, o resultado das operações incongruentes de trocas de palavras não interferiu na distribuição de categorias gramaticais. Tanto a distribuição de palavras entre gramaticais e lexicais, quanto a distribuição das diversas categorias lexicais, permaneceram inalteradas.

Essa estabilidade das categorias gramaticais, mesmo com a troca incongruente de palavras, é um fenômeno cujas possíveis origens merecem atenção. O que podemos imaginar é que as trocas de palavras são incongruências segmentares; que ocorrem pontualmente ao longo do fluxo linguístico;

que sofrem influência do contexto semântico e fonológico onde ocorrem (como discutiremos a seguir); mas que também parecem sofrer influência de um contexto gramatical, estrutural, que limita as escolhas dos leitores quanto às palavras a serem inseridas.

Com relação às palavras gramaticais, faz sentido pensar que elas sejam indispensáveis à organização sintática dos períodos e, portanto, dificilmente sofram mudança de categoria nas trocas (sofreram apenas 3% de trocas de categoria). Da mesma forma, quanto às palavras lexicais, também faz sentido pensar que suas categorias gramaticais (substantivo, adjetivo, advérbio, verbo) estejam de certa forma "amarradas" ao contexto sintático (e semântico) onde ocorrem. Dessa forma, a vizinhança linguística de um verbo não propiciaria sua substituição inadvertida por um substantivo, por exemplo; e assim por diante. E, de fato, os verbos sofreram troca de categoria em apenas 15% dos casos, e os substantivos, em apenas 20% dos casos. Em contrapartida, adjetivos, participípios (que, frequentemente, exercem função adjetival) e advérbios mudaram de categoria em mais de metade dos eventos de troca de palavra. Essa é uma manifestação linguística concreta de uma provável diferença na natureza dessas palavras (verbos e substantivos, de um lado, adjetivos, advérbios e participípios, de outro), na forma como elas estão cognitivamente registradas na mente dos falantes.

5.3 As incongruências suprasegmentares

Como já mencionamos, as incongruências suprasegmentares têm natureza diferente das incongruências segmentares. Estas deformam as palavras e aquelas, o fluxo normal da fala e/ou sua segmentação. As incongruências suprasegmentares, por si sós, talvez também merecessem uma cisão. Pausas, hesitações e lentificações do fluxo da fala, por um lado, e incongruências de fronteira, por outro, talvez devessem ser interpretadas como fenômenos (bastante) distintos. O primeiro grupo parece revelar uma dificuldade de processamento de aspectos da linguagem, do fluxo serial das palavras, considerando a velocidade com que vinham sendo processados até o momento. Pausas, hesitações e lentificações refletiriam, então, dificuldades de processamento que estariam materializadas como redução na taxa de elocução; e, por retardar determinadas decisões, poderiam até mesmo evitar o surgimento de outras incongruências. O segundo grupo, o das incongruências de fronteira, por outro lado, parece revelar que o falante já tomou decisões com relação à organização do período e que tais decisões não corresponderam à realidade linguística pretendida pelo texto escrito. Um grupo adicional, heterogêneo, chamado, em nosso trabalho de "outros", agrupou incongruências que não fomos capazes de classificar e que, via de regra, incluíram aspectos de entoação que pareceram incongruentes, mas

aos quais não conseguimos atribuir categorias; foram, então, registrados como incongruências *lato sensu* e não foram analisadas mais a fundo.

Duas outras incongruências que, a priori, imaginamos que poderiam ser relativamente comuns nos surpreenderam pela quase inexistência. Pensamos que haveria um bom número de casos de períodos interrogativos lidos como declarativos e vice-versa. Pensamos, até mesmo, que a adoção do ponto de interrogação invertido (¿) usado pela língua espanhola no início de frases estaria plenamente justificada pelos resultados que obteríamos em nosso estudo. ¿Estávamos corretos? Não: registramos apenas 3 ocorrências do tipo, ou 0,3% de todos os períodos do corpus. Ou seja, não se trata de uma questão relevante em PB.

Lentificações do fluxo da fala também não foram comuns: identificamos apenas 17 eventos, correspondendo a 2% das incongruências suprasegmentares, ou 0,8% de todas as incongruências. As lentificações, quando não são usadas voluntariamente como recurso expressivo, retórico, devem indicar dificuldades de processamento. O que distingue a lentificação da pausa é que a primeira ocorre durante a emissão sonora e a segunda, entre emissões; entre palavras. A predominância absoluta das pausas sobre as lentificações (294 pausas vs. 17 lentificações) deve ter alguma justificativa do ponto de vista do processamento da linguagem. Pode-se imaginar que, pelo menos na fala lida, as palavras sejam unidades linguísticas cuja emissão sonora, uma vez iniciada, tenda ser completada. Isso deve ser verdade, principalmente, para palavras gráficas que estejam cognitivamente bem registradas no léxico mental do falante, como um *chunk*, um aglutinado, uma unidade cognitiva autônoma que, sendo identificada no texto, é emitida em um único gesto articulatório.

As pausas foram, então, a incongruência suprasegmentar individual mais comum em nosso corpus, com 34% do total de incongruências, quase 300 eventos, ou cerca de 1 pausa por minuto de fala, em média. Como mencionamos acima, as pausas são reflexo de alguma dificuldade de processamento da linguagem e mereceriam um estudo mais detalhado. Poderíamos ter correlacionado suas ocorrências com diversas variáveis, como outras incongruências, por exemplo. Por razões de planejamento e cronograma, não fizemos essa análise.

5.3.1 As incongruências de fronteiras prosódicas

As inserções de fronteira não terminal foram a segunda incongruência suprasegmentar mais comum, totalizando 28% dessas incongruências. Se considerarmos os quatro tipos de incongruências de fronteira como uma categoria à parte, como propusemos, essa categoria – incongruências de

fronteira — totalizaria 52% de todas as incongruências suprasegmentares. E se considerarmos exclusivamente as inserções de fronteira (e não as deleções), elas totalizariam 42% de todas as incongruências suprasegmentares, superando até mesmo as pausas. Esses números mostram que as incongruências de fronteiras prosódicas foram bastante comuns em nosso corpus, acumulando 448 ocorrências, cerca de 1,5 por minuto de fala, ou quase uma a cada 3 períodos, em média.

5.3.1.1 Por que as incongruências de fronteira não são objeto de mais estudos?

As incongruências de fronteira, conforme as definimos, não têm sido muito investigadas na literatura; talvez por não serem comuns na fala espontânea (se é que ocorrem nela). Trata-se de uma questão interessante, porque a dinâmica da fala espontânea — em que o falante produz a fala a partir de uma ideia e, para avaliar a felicidade do que produziu, monitora continuamente sua produção sonora — pode ser, de certa forma, incompatível com a produção de incongruências de fronteira ou, pelo menos, com a permanência da incongruência sem que haja retratação. Por outro lado, Scarpa (2012, p. 28), analisando a estrutura prosódica de disfluências em PB, diz que "a fluência é uma abstração metodológica, baseada na leitura ensaiada ou 'profissional' de um texto escrito ou em textos orais decorados e ensaiados". Então, se não observamos incongruências de fronteira na fala espontânea (ao que me parece), e se "fluência é uma abstração (...) baseada na leitura ensaiada", compreende-se porque estudos de fala lida costumem ignorar falantes adultos, escolarizados, sem patologias de fala. Aparentemente, o produto desse tipo de leitura estaria muito próximo do modelo estabelecido de "fluência" (congruência com o texto). Com base nisso, nossa referência para incongruências de fronteira buscou comportamentos de segmentação prosódica que não esperaríamos encontrar em um modelo de congruência, como apontou Scarpa (2012), nem na fala espontânea, como a conhecemos.

5.3.1.2 Como foram validadas

As incongruências de fronteira, em nosso estudo, foram identificadas por critérios perceptivos, e esses critérios foram validados, como descrevemos, por 4 falantes de PB. A concordância quanto à existência de incongruências de fronteira, segundo o teste Kappa de Cohen, variou de moderada a substancial. O voluntário escolhido para validar, de forma mais precisa, a marcação específica do local de ocorrência e do tipo de incongruência concordou com o pesquisador, nessa análise mais detalhada, em grau moderado. Esses resultados indicam que a percepção da existência de incongruências de fronteira não é um artifício criado pelo pesquisador, já que são percebidas também por falantes de PB

que nunca tiveram contato metacognitivo prévio com a linguística, descontados os estudos da gramática normativa obrigatórios na educação básica brasileira.

5.3.1.3 Por que as incongruências suprasegmentares não são retratadas?

Um achado inesperado com relação às incongruências suprasegmentares é a ausência quase total de retratações. Em um universo de 858 eventos incongruentes, identificamos apenas 4 retratações (0,5%), contra 305 (26%) retratações no universo das incongruências segmentares. Isso não poderia escapar de uma análise e de, pelo menos, algumas hipóteses. O primeiro passo para essa análise foi separar, como propusemos, as incongruências em dois grupos. O primeiro grupo foi o das pausas, hesitações e lentificações, com 362 eventos. A ausência absoluta de retratações nesse primeiro grupo de incongruências deve associar-se ao fato de que elas não são, propriamente, erros. O leitor não falou algo que não estava escrito, apenas retardou a continuidade da leitura, sem emitir sons adicionais e sem trocar os correspondentes fonológicos do texto escrito por outros (note-se que não identificamos em nosso corpus os tais “*fillers*”, hesitações entre palavras). Dito isso, falar “irretocáááá...vel”, hesitantemente, é diferente de dizer “irrecotável”. “Irrecotável” é um desvio de ortoépia que, provavelmente, induziria uma retratação. É possível, também, que o leitor esteja atento à tarefa autoimposta de passar uma mensagem e, por isso, opte voluntariamente por não reler o trecho em que houve pausa, hesitação ou lentificação, julgando que a releitura não seria produtiva. Se a tarefa fosse ler o texto com a maior perfeição possível, com aderência absoluta ao que está escrito, talvez observássemos retratações desses tipos de incongruência (pausas, hesitações e lentificações) com frequência maior do que observamos. Essa é uma dúvida que permanecerá e que só será sanada, creio, por meio de um estudo experimental cuja variável independente seja a tarefa imposta aos participantes e a variável dependente, a incidência de retratações.

O segundo grupo de incongruências é o das incongruências de fronteira. Ignoraremos o conjunto das incongruências classificadas como “outros”, apesar de serem, predominantemente, incongruências de entoação que, se não puderam ser categorizadas como incongruências de fronteira, não deixam de ser fenômenos prosodicamente assemelhados. O grupo das incongruências de fronteira, com 448 ocorrências, resultou em apenas 4 retratações, ou menos de 1%. Não podemos dizer que este grupo equivale ao anterior (pausas, hesitações e lentificações) quanto à interpretação de que não se trata de erros. As incongruências de fronteira podem ser consideradas como um desvio do que seria a forma prosódica minimamente adequada ao conjunto de sintagmas de um enunciado, com suas unidades acentuais e entoacionais. É um desvio que ocorre tanto quando há omissão de fronteiras —

o que ocorreu em 20% dos casos —, como quando há inserção de fronteiras — o que ocorreu nos restantes 80% das vezes.

No entanto, mesmo que nossa posição seja a de considerá-las como "erros", parece que os falantes não concordam conosco (pelo menos, não tanto quanto com relação às incongruências segmentares); ou, simplesmente, não conseguem perceber essas incongruências como nós as percebemos. O fato é que os falantes não se retratam quando produzem uma incongruência de fronteira. E por que não se retratam? Nossas hipóteses, não mutuamente excludentes, são as seguintes:

- Os falantes estão mais preocupados em ler do que em compreender o texto e, portanto, não costumam perceber as incongruências prosódicas que produziram
- A percepção da incongruência prosódica de fronteira depende de uma percepção global do enunciado e de frases intermediárias, como grupos acentuais (que consideraremos, como alguns autores, sinônimo de palavra fonológica) e grupos entoacionais; sendo assim, o falante não percebe imediatamente a incongruência e quando (ou se) a percebe, já não é conveniente retratar-se, pois a retratação envolveria retornar a um ponto anterior da leitura e repetir um segmento relativamente grande do enunciado
- A percepção do autor e dos voluntários com relação às incongruências de fronteira prosódica, ao contrário dos falantes participantes do estudo, não foi uma percepção automática, espontânea, improvisada. Os juízes foram preparados para encontrá-las; leram os períodos escritos antes de ouvir os enunciados falados pelos leitores; estavam atentos, "sedentos" por encontrar incongruências, sem outra tarefa paralela que lhes ocupasse o espaço atencional. Portanto, é bem possível que, numa tarefa de leitura em voz alta, a capacidade de perceber eventuais incongruências de fronteira seja radicalmente menor do que em uma tarefa que envolva, pura e simplesmente, encontrá-las

Essa última hipótese deve ser considerada com atenção. Ela nasce da própria motivação para a realização deste trabalho. O contato inicial do autor com as incongruências de fronteira partiu de sua percepção de que textos conhecidos por ele foram lidos em voz alta de uma forma que não correspondia ao que estava escrito. Nessas situações, mesmo notando todos os tipos de incongruências (segmentares e suprasegmentares) que, na fala lida, distorceram o texto escrito, as incongruências de fronteira eram especialmente perceptíveis, pois desestruturavam toda a organização sintagmática dos enunciados. Essas incongruências demonstravam, de forma bastante evidente, que o leitor não apenas não estava

falando espontaneamente, como estava lendo "errado". Portanto, é possível que um conhecimento prévio do que vai ser lido seja um pré-requisito importante para a percepção de incongruências de fronteira na fala lida. E um experimento interessante para validar esse "pré-requisito" seria empregar o conhecimento prévio do trecho a ser lido (conhecido vs. não conhecido) como variável independente, e a percepção, ou não, de incongruências de fronteira como variável dependente.

5.3.1.4 Incongruências de fronteira prosódica: correlação com a segmentação do discurso

Com as considerações anteriores em mente, analisamos a correlação entre incongruências de fronteira e duas variáveis linguísticas: a segmentação do discurso, adaptado de Grosz e Sidner (1986) e discutido nesta subseção, e a estratégia de identificação do ator, adaptado de Bornkessel-Schlesewsky e Schlewsky (2014) e discutido na subseção seguinte.

Nos capítulos Revisão da Literatura e Metodologia, ressaltamos a subjetividade dos critérios utilizados por Grosz e Sidner (1986) para a segmentação linguística do discurso. Tal subjetividade vem-se tornando evidente também para outros pesquisadores, como declarou recentemente a professora Luciana Lucente durante a sessão virtual de defesa de dissertação de mestrado de Jesse da Silva Lima, na Universidade de Brasília. Procuramos, então, encontrar um critério mais objetivo para essa análise e optamos pela continuidade/descontinuidade de tópico ou foco entre dois períodos adjacentes.

O tema tópico/foco, com nomenclaturas e abordagens diversas, não é algo pacificado, como bem aponta Partee (1991). Nossa abordagem, entretanto, é menos sujeita a divergências, porque não diferencia tópico de foco, considerando-os em conjunto. Como esclarecemos no capítulo Metodologia, nosso critério para descontinuidade discursiva (DS1) do período atual com relação ao período anterior foi que o tópico do período atual não tivesse sido mencionado no período anterior, seja como tópico ou como foco. Isto é, o período atual, que está sendo analisado quanto à presença ou não de incongruências, fala sobre algum assunto que não foi mencionado no período imediatamente anterior.

Analisada dessa forma, a continuidade, a continuidade temática associou-se à presença de incongruências de fronteira prosódica. Períodos cujos tópicos haviam sido mencionados no período anterior apresentaram uma maior incidência de inserção/deleção incongruente de fronteiras prosódicas. Isso é, de certa forma, surpreendente. Apesar de não termos levantado hipóteses prévias com relação ao que encontraríamos, teria sido sensato imaginar que poderia haver associação entre previsibilidade e congruência. O problema é que a continuidade temática, aparentemente, não aumenta a

previsibilidade, mas apenas a familiaridade. Períodos tematicamente contínuos são mais familiares aos leitores, mas não são mais previsíveis. É bem possível que a familiaridade gere no leitor uma falsa sensação de segurança, de previsibilidade.

A hipótese de que a familiaridade possa ser nociva quando se trata da leitura em voz alta resgata dois conceitos relacionados, que especificam melhor a noção genérica de “familiaridade”. O primeiro é o de predictive processing (CHRISTIANSEN; CHATER, 2015, p. 9), um mecanismo ativo que “explora múltiplas restrições em paralelo, em diferentes níveis de representação linguística”¹. O segundo é o de preparedness (FERREIRA; CHANTAVARIN, 2018, p. 445), que sugere que “uma informação dada faz contato com o material linguístico e o conhecimento prévios, e a integração do input atual com o contexto e o conhecimento precedentes criam uma rica representação semântica”². Pois bem, parece que, na leitura, nada disso ajuda. Ou, na verdade, até atrapalha.

Um terceiro conceito que faz contato com a ideia genérica de “familiaridade” é o de priming (CAPLAN, 2016). De acordo com esse fenômeno, o período atual é sempre enviesado com relação a características sintáticas, semânticas, ou mesmo prosódicas de períodos anteriores. É possível que uma conexão semântica com o período imediatamente anterior reforce a expressão de características latentes que sofreram priming e, assim, prejudique a expressão das reais características do período que está sendo lido.

5.3.1.5 Incongruências de fronteira prosódica: correlação com a estratégia de identificação do ator

Outra análise que fizemos abordou a correlação entre incongruências prosódicas de fronteira e a estratégia de identificação do ator, conforme Bornkessel-Schlesewsky e Schlewsky (2014). Uma primeira avaliação — que incluiu o conjunto de todas as categorias analisadas e discriminou os sujeitos em 38 diferentes tipos — demonstrou haver correlação entre classificação do sujeito (em graus de agentividade) e incongruências de fronteira. Na sequência, para identificar qual subtipo classificatório

¹ “We have already noted that, to be able to chunk incoming information as fast and as accurately as possible, the language system exploits multiple constraints in parallel across the different levels of linguistic representation”.

² “Information that is given makes contact with linguistic material that came before, as well as with background knowledge, and integration of the input with preceding context and knowledge leads to the creation of a rich semantic representation.”

poderia estar associado a congruência ou incongruência, analisamos cada subcategoria isoladamente — voz ativa/passiva, sujeito desinencial, animacidade, definitude, pessoa e posição. Um detalhe que vale mencionar é que não conseguimos avaliar se há diferença entre sujeitos humanos e não humanos, pois descobrimos, durante a análise, que nosso corpus não incluiu sujeitos animados não humanos (animais).

Dito isso, não encontramos correlação entre nenhuma das características analisadas e congruência/incongruência, exceto no caso de sujeitos em primeira pessoa. Neste caso, existe uma associação significativa entre períodos com sujeito em primeira pessoa e ausência de incongruências de fronteira; mas esses sujeitos têm duas características adicionais que precisam ser mencionadas. Praticamente todos os sujeitos em primeira pessoa estavam em posição pré-verbal, e 88% deles eram sujeitos desinenciais marcados na flexão morfológica do verbo (p.ex.: "fi-lo porque qui-lo"). Ajustando os resultados a essas observações, podemos afirmar que períodos com sujeito em primeira pessoa — que, em nosso corpus, ocorreram na posição pré-verbal e eram, predominantemente, desinenciais — associaram-se a menor incidência de incongruências de fronteira prosódica.

Restou, entretanto, uma dúvida. Apesar de termos visto que sujeitos desinenciais de um modo geral não se correlacionam com congruência/incongruência de fronteiras pródicadas na leitura, é possível que haja alguma correlação quando se trata especificamente de sujeitos em primeira pessoa. Para afastar a dúvida, aplicamos o *Pearson's Chi-squared test with Yates' continuity correction* a uma tabela comparando sujeitos em primeira pessoa desinenciais vs. não desinenciais, e congruência vs. incongruência, o obtivemos um *p-value* de 0.2502. Em suma, primeira pessoa associa-se a congruência, independentemente de haver sujeito desinencial ou expresso.

Levantamos duas hipóteses para explicar esse achado, sendo que uma não exclui a outra. A primeira é uma hipótese semântica. A ideia é que frases, orações e períodos em primeira pessoa limitariam o alcance daquilo que pode ser dito, pois todos os fatos relacionados a esse sujeito são conhecidos e restritos ao universo pessoal. Quando se fala de segundas ou terceiras pessoas, as possibilidades são ilimitadas e a imprevisibilidade é muito maior. Acreditamos que o conhecimento sobre o que pode ser dito limita as possibilidades do que vai ser dito e pode facilitar a organização prosódica do período.

A segunda hipótese é uma hipótese gramatical. O sistema pronominal do PB — mesmo considerando o sistema em uso (MENON, 1995), em oposição, e também em complemento, ao sistema

normativo, padrão — discrimina muito bem, em primeira pessoa, o agente (a) do paciente (b), sendo a) o agente do verbo transitivo ou argumento único do verbo intransitivo (que seriam sempre "eu"); e b) o paciente do verbo transitivo (que seria sempre "me" ou "mim"). Essa diferenciação, que é um tipo de marcação de caso, não está sempre presente em segunda ou em terceira pessoa. Assim, um sujeito sintático em primeira pessoa é facilmente diferenciado de um objeto sintático em primeira pessoa e isso deve contribuir para a percepção do leitor quanto à correta organização sintagmática e prosódica do período. Nos termos do modelo em que nos inspiramos para fazer essa análise (BORNKESSEL-SCHLESEWSKY E SCHLESEWSKY, 2014), a "identificação do ator" e dos demais participantes da estrutura argumental do verbo torna-se muito mais simples e automática quando o sujeito (em PB) está em primeira pessoa.

Outro aspecto dessa hipótese gramatical relaciona-se ao fato de que a primeira pessoa é sempre pronominal, ao contrário da terceira pessoa (desconsideraremos, nessa afirmação, situações como "dê o copo para o papai aqui" ou "Os brasileiros não sabemos votar"). A terceira pessoa, além de ser muito frequentemente não pronominal ("o cão mordeu a capivara"), quando o é, não diferencia bem, no sistema pronominal atualmente em uso ("ele mordeu ela"), os casos nominativo e acusativo ("ele a mordeu"). Quanto à segunda pessoa, apesar de também ser pronominal, não diferencia muito bem (também considerando o sistema pronominal em uso) os casos gramaticais. Hoje em dia, são usuais expressões do tipo "dei o doce para você", em vez de "dei-lhe o doce", ou "dei o doce para ti". Ou seja, o "você", como pronome pessoal de segunda pessoa, não costuma ser marcado para caso. Além disso, em nosso corpus, por suas próprias características, a segunda pessoa teve uso bem pouco frequente.

5.4 Predicação primária e hierarquia argumental

Essa análise envolveu conceitos não tão bem delimitados na literatura e uma metodologia também pouco usual. Além disso, por se tratar de uma decisão *post hoc*, não houve tempo hábil para validar nossas percepções com voluntários falantes de PB, como fizemos com as inserções/deleções de fronteiras prosódicas. De qualquer forma, julgamos que seria importante fazer a análise, mesmo sendo necessário ressaltar suas possíveis limitações.

Ao longo do estudo, quando observamos que as incongruências de fronteiras prosódicas, apesar de serem bastante frequentes, não se correlacionavam com algumas variáveis linguísticas que escolhemos analisar (taxa de elocução, extensão dos períodos, estratégia de identificação do ator) e que, além disso, não eram retratadas, passamos a refletir mais detidamente sobre o assunto e a

investigar o que a literatura tinha a dizer a respeito. Parecia necessário buscar algum fator mais específico, alguma variável linguística que eventualmente pudesse se correlacionar com as incongruências de fronteiras prosódicas.

Não encontramos nenhuma abordagem concreta, única, que explicasse nossos (não) achados, mas encontramos diversos estudos e teorias que tangenciavam a questão e que poderiam ser aplicados ao nosso problema. Como comentamos no capítulo Metodologia, optamos por utilizar os conceitos propostos por Isre'el (2018) para segmentar cada período dos textos escritos em "enunciados" potenciais (segmentos com autonomia pragmática e interpretabilidade em isolamento, a que denominamos, em casos específicos, "predicação primária") e "módulos de informação" em potencial (segmentos que, associados à predicação primária, não necessitariam de nenhum material linguístico adicional para fazer sentido).

Nossa hipótese era que a correspondência apontada por Isre'el (2018) entre fronteiras prosódicas e unidades discursivas poderia explicar, ao menos em parte, as incongruências que observamos. Sob essa perspectiva, a reconstrução oral, pelos leitores, da estrutura prosódica (e discursiva) dos períodos escritos esbarraria nas diversas possibilidades oferecidas por cada nova palavra surgida ao longo da sequência de palavras que, ao fim, completariam o período. Como o leitor, em sua reconstrução oral, não parte do período completo e não tem uma ideia global do enunciado, resta-lhe apenas ir levantando a parede prosódica, tijolo por tijolo, palavra por palavra, sem dispor de uma planta detalhada que lhe oriente a construção. Com base nessa hipótese geral, levantamos algumas hipóteses específicas para os resultados que esperávamos obter a partir da segmentação dos períodos segundo sua "hierarquia argumental":

- a) As predicações primárias estariam preservadas (não seriam fragmentadas)
- b) Não haveria inserção de fronteira terminal antes de completar a predicação primária
- c) Haveria mais incongruências de fronteira depois da predicação primária
- d) As inserções incongruentes de fronteira ocorreriam, majoritariamente, nos desnivelamentos (quebras de linha) da hierarquia argumental

A hipótese "a", de certa forma, se confirmou, mas em combinação com a hipótese "d". Nossa justificativa para a hipótese "a" era de que potenciais enunciados autônomos contidos dentro dos

períodos seriam resistentes à segmentação indevida, pois suas palavras guardariam, entre si, uma força de coesão semântica e sintática maior do que as demais palavras do período. O que observamos, na verdade, foi que a congruência não estava propriamente associada à relação entre palavras adjacentes pertencentes à predicação primária, mas a quaisquer palavras adjacentes cujas relações hierárquicas fossem niveladas. Ou seja, houve inserção/deleção de fronteiras prosódicas entre palavras pertencentes à predicação primária, mas com incidência semelhante à que observamos entre palavras com quaisquer tipos de relações hierárquicas niveladas. Por outro lado, tais inserções/deleções de fronteiras ocorreram em quantidade muito maior entre palavras com relações desniveladas, levando-nos a concluir que não foi o fato de compor a predicação primária que evitou a inserção/deleção incongruente de fronteiras prosódicas, mas, sim, o fato de haver, entre palavras adjacentes, uma relação nivelada.

O conceito de "relação nivelada" merece um aprofundamento. Como comentamos no capítulo Metodologia, a classificação das palavras de um período em níveis hierárquicos remete à ideia de argumentos e adjuntos. Uma relação hierárquica nivelada entre duas palavras indica que a segunda palavra é exigida, é necessária, é obrigatória para a complementação semântica e/ou sintática da primeira palavra. Na frase "Essa corporação, cujo aniversário de...", a palavra "corporação" (ou alguma outra, equivalente) não pode ser dispensada após a produção (produção via leitura, no nosso caso) de "essa" e, portanto, está em relação nivelada com ela. A palavra "essa" não se satisfaz sozinha, "exigindo" algum tipo de complemento. O mesmo ocorre com "aniversário", em sua relação com "cujo". A ideia de nivelamento, então, é análoga, mas não idêntica, à ideia de argumento, em que uma palavra é argumento de outra quando é exigida para complementá-la.

Em nossa abordagem, utilizamos um critério objetivo para definir relações niveladas ou desniveladas: a obrigatoriedade. Sempre que determinada palavra podia ser omitida sem haver comprometimento da gramaticalidade do período e, adicionalmente, de sua coerência semântica e pragmática com o discurso proferido até aquele momento (o que, de certa forma, equivale a dizer que a omissão da palavra não afetaria a "felicidade", num sentido austiniano, do período), sempre que uma palavra preenchia tais critérios de omissibilidade, ela era deslocada para um nível inferior na hierarquia. Dito de outro modo, a existência dessa palavra — naquele ponto específico do período, sucedendo a palavra anterior — não era obrigatória. No entanto, nas discussões da literatura sobre a realidade linguística de argumentos, adjuntos e abstrações correlatas, o critério obrigatoriedade não costuma ser uma fronteira absoluta a separar argumentos de adjuntos.

Merlo e Ferrer (2006, p. 343) reconhecem a dificuldade de classificar itens lexicais em argumentos vs. adjuntos, acrescentando que "*researchers have abstracted away from this distinction, because identifying arguments and adjuncts is a notoriously difficult task, taxing many native speakers' intuitions*"¹. Uma das dificuldades, segundo as autoras, é que, apesar de os adjuntos serem sempre opcionais, os argumentos não seriam sempre obrigatórios. Sob esse ponto de vista, a justaposição do conjunto dos argumentos com o conjunto dos adjuntos formaria um diagrama de Venn cuja intersecção abrigaria os argumentos considerados "não obrigatórios". Daí o emprego usual de critérios adicionais e não homogêneos de classificação, além da obrigatoriedade/opcionalidade, para separar os tais "argumentos não obrigatórios" dos adjuntos.

A noção de argumento associa-se ao conceito de estrutura argumental. A estrutura argumental de um determinado item lexical representaria outros itens lexicais (argumentos) cujas presenças seriam exigidas e determinadas pelo primeiro item, considerado o titular da estrutura argumental. E, a partir desse conceito, surge a ideia de "valência", que representa o número de argumentos que compõem a estrutura argumental de determinado item lexical, especialmente quando se trata de verbos. No entanto, apesar de se referir, usualmente, aos verbos, o conceito também é aplicado a outras categorias gramaticais, como nomes, adjetivos e até preposições (LEVIN, 2013). No caso dos verbos, a estrutura argumental, a depender do autor e da teoria, recebe nomes diversos, como complementação, subcategorização, modelo de regência (*government model*), ou *sautzbauplan* (*clause blueprint*, um desenho esquemático da oração), entre outros; e a própria distinção argumento/adjunto também já recebeu outros nomes, como complemento/modificador, argumento/satélite e actante/circunstante (HASPELMATH; HARTMANN, 2015).

O próprio Haspelmath (2014), entretanto, reconhece que as definições atuais de argumento e adjunto provavelmente não se aplicam translinguisticamente, e os critérios utilizados por cada língua para essa distinção acaba criando categorias linguísticas não universais. Essa posição é endossada por Creissels (2014), mas com a ressalva de que o uso de critérios semânticos posicionaria os participantes de qualquer evento representado por um verbo em uma escala contínua, não dicotômica, que iria de participantes com características prototípicas de argumentos a participantes com características prototípicas de adjuntos. Além disso, ele divide os "argumentos não obrigatórios" em dois grupos. No

¹ Os pesquisadores têm evitado essa distinção, porque a identificação de argumentos e adjuntos tem sido uma tarefa reconhecidamente difícil, confundindo a intuição de muitos falantes nativos

primeiro grupo, estariam os argumentos que poderiam ser omitidos por já terem sido mencionados anteriormente ou por estarem evidentes pelo contexto; ou seja, participantes que estariam, de alguma forma, presentes, e não seriam realmente omitidos. O segundo grupo incluiria argumentos ausentes, não especificados, não evidentes pelo contexto, nem previamente mencionados. Ou seja, seriam argumentos com características muito semelhantes às de adjuntos. Em uma hipotética escala contínua entre argumentos e adjuntos, esses participantes estariam, talvez, no centro da escala, um pouco deslocados para o polo dos argumentos.

Em contraste com essas propostas, nossa abordagem, em nome da objetividade, considerou apenas o critério da obrigatoriedade. Cada nova palavra, no momento em que surgia, considerando apenas as palavras surgidas até aquele momento e ignorando as palavras ainda não lidas, era classificada como obrigatória (não omissível) com relação às palavras anteriores, ou não obrigatória (omissível) com relação às palavras anteriores. E a relação de obrigatoriedade foi hierarquizada, pois a palavra em análise poderia ser obrigatória com relação à palavra imediatamente anterior, estando em uma relação nivelada com ela, mas não ser obrigatória com relação a outros níveis hierárquicos superiores. O exemplo a seguir relembra os critérios que utilizamos:

- Nas relações niveladas, as palavras não são omissíveis; e, no exemplo abaixo, as três primeiras palavras, além de estarem no mesmo nível hierárquico (niveladas), compõem o primeiro enunciado em potencial do período, correspondendo a sua predicação primária:
 - José esfaqueou João...
- O período, entretanto, não termina em "João", e o leitor tem de estar atento ao fato de que a ideia completa pretendida pelo período exige a inclusão de todas as suas palavras; mas é possível, semântica e sintaticamente (caso não se saiba que o período deveria prosseguir), interrompê-lo em diversos momentos, como ilustrado, sequencialmente, abaixo:
 - José esfaqueou João e Juliana...
 - José esfaqueou João e Juliana, que correu...
 - José esfaqueou João e Juliana, que correu com o sorvete...
 - José esfaqueou João e Juliana, que correu com o sorvete na mão...
 - José esfaqueou João e Juliana, que correu com o sorvete na mão, derrubando a rosa...
 - José esfaqueou João e Juliana, que correu com o sorvete na mão, derrubando a rosa no chão...
 - José esfaqueou João e Juliana, que correu com o sorvete na mão, derrubando a rosa no chão do parque...

- José esfaqueou João e Juliana, que correu com o sorvete na mão, derrubando a rosa no chão do parque, perto da Boca do Rio.

Cada nova linha acima ilustra a adição, ao período, de novas palavras que guardam, entre si, relações hierárquicas niveladas (p.ex.: "que correu", "com o sorvete", "na mão" etc). O período completo, o enunciado pretendido pelo texto escrito, entretanto, só está estará satisfeito no último item da enumeração. A ideia que nos levou a optar por esse tipo de segmentação foi imaginar que, a cada nova linha, quando há a inserção de material linguístico "desnivelado" com relação à linha superior, o leitor pode ser sido induzido a interromper, indevidamente, o período e inserir uma fronteira prosódica incongruente.

Outro aspecto que nos afastou da abordagem "tradicional" sobre a distinção argumento vs. adjunto foi o fato de que a literatura analisa o material linguístico a posteriori, considerando os períodos retrospectivamente. Em nossa análise, tal abordagem pareceu contraintuitiva, pois os leitores interpretam os períodos prospectivamente, a cada nova palavra, em tempo real, sem saber exatamente o que vem pela frente. E eles tomam decisões sobre a estruturação prosódica do período sem ter o luxo de conhecer, como fazem os linguistas, o período completo. Um exemplo da abordagem linguística tradicional, proposta por Haspelmath (2014), é a seguinte:

1. *I wrote a letter.* > **I wrote, and I did a letter*¹
2. *I wrote with a pen.* > *I wrote, and I did it with a pen*

Em (1), estaria comprovada a condição de argumento de "*a letter*", pois sua existência torna-se possível por intermédio do verbo *wrote*; e, assim, "*a letter*" não poderia ocorrer livremente, com um verbo qualquer, como, por exemplo, o verbo *to do*. Em (2), ao contrário, "*a pen*" seria apenas adjunto, e poderia ocorrer com outro verbo qualquer, sem incorrer em agramaticalidade. Nossa crítica a essa abordagem, na situação específica do nosso estudo, relaciona-se ao fato de o leitor chegar a *wrote* sem saber o que vem depois: se virá um argumento, um adjunto ou, até mesmo, se não virá absolutamente nada. Em PB, seria perfeitamente aceitável (sintaticamente e, talvez, a depender do contexto, semanticamente) interromper o período na altura de *wrote*, caso não se percebesse que ainda havia material linguístico a ser anexado ao verbo, como, por exemplo, em "eu escrevi (uma carta)(à caneta)".

¹ (1) Eu escrevi uma carta. > * Eu escrevi, e eu o fiz a uma carta. (2) Eu escrevi com uma caneta. > Eu escrevi, e eu o fiz com uma caneta. ([tentativa de] tradução nossa)

No nosso estudo, o material linguístico entre parêntesis nesse exemplo foi considerado apenas como um elemento omissível, independentemente de poder ser classificado, por critérios heterogêneos e variados, como adjunto ou argumento.

Quanto às hipóteses "b" e "c", que julgavam que "não haveria inserção de fronteira terminal antes de completar a predicação primária" e que "haveria mais incongruências de fronteira depois da predicação primária", descobrimos, em primeiro lugar, que não se devem fazer afirmações categóricas como "b". Ao contrário do que afirmamos, houve inserção de todo tipo de fronteira antes da e durante a predicação primária, apesar de terem ocorrido em menor quantidade. E, quanto a "c", como não medimos a extensão, nem a posição relativa, das predicções primárias com relação aos períodos, não pudemos tirar conclusões absolutas sobre as incidências de cada tipo específico de incongruência de fronteira com relação à predicação primária, pois o achado de que, por exemplo, apenas 7% das inserções de fronteira terminal (IT) ocorreram durante as predicções primárias poderia indicar, simplesmente, que as predicções primárias ocupam apenas 7% da extensão dos períodos.

Apesar disso, pudemos comparar as incidências das incongruências de fronteira entre si, e observamos que as inserções de fronteira (em contraste com as deleções) ocorrem, predominantemente, após a predicação primária, e que essa tendência é mais forte com relação as IT. Esse resultado quanto às IT parece lógico, pois as IT indicam o final de um período/enunciado e, portanto, tenderiam a ocorrer em posições mais tardias do que as inserções de fronteira não terminal (INT).

É importante ressaltar que a predicação primária é, por definição, apenas o primeiro enunciado em potencial surgido ao longo do período. Em muitos casos — como, por exemplo, em orações coordenadas —, havia outro enunciado em potencial no mesmo período, mas esse enunciado não foi considerado como a predicação primária porque um enunciado anterior já havia ocupado essa posição. Enunciados em potencial surgidos após outros enunciados em potencial foram identificados, então, apenas como uma sequência de palavras em relação hierárquica nivelada, não sendo diferenciados de sequências de palavras (não verbais) de coocorrência obrigatória (p.ex.: 'pão de açúcar'). Isso indica que relações niveladas em posição hierárquica maior do que "1" identificam apenas conjuntos de palavras codependentes, incluindo, mas não se limitando a, verbos e seus argumentos obrigatórios.

Dito isso, observamos que as incongruências de fronteira se associaram a relações hierárquicas desniveladas, o que faz sentido quando resgatamos os critérios aplicados nessa classificação. Uma

relação desnivelada entre duas palavras indica que, na sequência da leitura, a segunda palavra não faz falta à estrutura sintática e/ou à composição semântica do período. Visto de outra forma, e resgatando a definição de argumento, a primeira palavra não "seleciona", não "pede", não "determina", não "exige" a segunda palavra; ou, indo além, a segunda palavra não faz parte da estrutura argumental da primeira palavra. Sendo assim, no momento em o leitor terminou o processamento da primeira palavra, mas ainda não iniciou o processamento da palavra seguinte, ele não tem nenhuma pista concreta (excluídas as pistas relacionadas à semântica do discurso até aquele momento) que indique qual será a relação da próxima palavra com a palavra que ele acabou de ler; e, em certos casos, ele não sabe nem mesmo se há outra palavra para ser incluída naquele período. Quando a relação é nivelada, por outro lado, a segunda palavra não é omissível, ela foi exigida e, até certo ponto, determinada pela palavra anterior e/ou pelo contexto discursivo e, portanto, o leitor, quando termina de ler a primeira palavra, já está, de certa forma, "avisado" de que alguma coisa será lida na sequência e, muitas vezes, já foi alertado sobre um leque limitado de possibilidades quanto ao que será lido. Sendo assim, não é surpreendente que relações niveladas inibam inserções/deleções incongruentes de fronteiras, e que relações desniveladas deixem em aberto todas as possibilidades para inserções/deleções de fronteira e, portanto, estejam associadas a elas.

Identificamos, também, associação entre incongruências de fronteira e nível hierárquico maior do que "2". Já havíamos observado que a predicação primária (relações niveladas de nível 1) e relações niveladas de um modo geral associavam-se a congruência, mas observamos também que períodos com uma proporção maior de níveis hierárquicos mais baixos (> 2) continham mais incongruências de fronteira. Infelizmente, não controlamos, nessa análise, a variável "extensão dos períodos". Períodos mais longos têm mais chance de ter níveis hierárquicos mais baixos e, apesar de termos observado anteriormente que a relação entre extensão de período e incidência de incongruências suprasegmentares é linear, não fizemos essa análise com relação, exclusivamente, às incongruências de fronteira. Não descartamos, portanto, que períodos mais longos, e não níveis hierárquicos mais baixos, associem-se a incongruências de fronteiras.

Outro fato que observamos foi que as inserções de fronteira são bem mais frequentes do que as deleções de fronteira. Ao todo, computamos 240 INT, 117 IT, 70 DNT e 21 DT, totalizando 357 inserções e apenas 91 deleções de fronteira. No cômputo geral, 80% das incongruências de fronteira foram inserções. A discussão do predomínio de inserções sobre deleções de fronteira pode ser feita de duas formas. A primeira analisaria aspectos autoevidentes dos fenômenos com que estamos lidando. A segunda buscaria enxergar o problema sob novas, possíveis, perspectivas. A explicação mais

autoevidente para a predominância das inserções de fronteira seria a de que qualquer justaposição de palavras poderia ser acompanhada por uma fronteira prosódica e essa fronteira seria incongruente sempre que, em situação normal e no contexto das demais fronteiras, não estivesse justificada. Se considerarmos que um enunciado contém a) no mínimo, uma fronteira terminal, e pode conter zero, uma, ou mais fronteiras não terminais, mas que b) via de regra, contém muito mais justaposições de palavras do que fronteiras prosódicas; compreendemos que há mais oportunidades para inserir fronteiras não existentes do que para deletar fronteiras existentes. A segunda hipótese, menos óbvia, será discutida a seguir.

5.5 Discussões adicionais

De início, discutiremos a segunda hipótese mencionada acima para tentar explicar o predomínio das inserções sobre as deleções de fronteira e, na sequência, acrescentaremos um adendo a este capítulo, abordando de forma mais livre o conjunto dos achados de nosso estudo e alguns modelos propostos pela literatura para explicar o processamento da linguagem. A ideia geral é que a predominância da inserção de fronteiras pode indicar uma tendência cognitiva para encerrar constituintes, em vez de prolongá-los. A inserção de uma fronteira organizaria o material linguístico produzido desde a fronteira anterior, e essa organização refletiria uma necessidade cognitiva básica. Partindo desse pressuposto, haveria uma tendência universal para inserir fronteiras, e não para eliminá-las. E, considerando essa justificativa hipotética para o predomínio das inserções sobre as deleções de fronteira prosódica e para a alta prevalência das inserções de fronteiras dentre todas as incongruências de leitura, discutimos brevemente a diferença entre fronteira terminal (FT) e não terminal (FNT), e entre inserção e deleção de fronteiras. Apesar de se tratar de uma discussão conceitual, que talvez estivesse melhor situada em capítulos anteriores, como revisão de literatura ou metodologia, optamos por inseri-la aqui porque estão incluídas certas conjecturas que derivaram de reflexões surgidas durante a análise dos resultados do trabalho.

5.5.1 A natureza das fronteiras prosódicas

Já mencionamos que aderimos à proposta de Isre'el (2018), segundo a qual as FTs indicariam uma transição entre enunciados e as FNTs indicariam transições entre módulos de informação. Do ponto de vista dos parâmetros fonético-acústicos, Raso, Teixeira e Barbosa (2020) sugerem que as FTs são mais bem definidas do que as FNTs. As FTs estão mais comumente associadas a pausas e discontinuidades da frequência fundamental, enquanto nas FNTs essas e outras variáveis fonético-

acústicas são muito menos padronizadas. Os autores sugerem, inclusive, que as FTs podem ter natureza diferente das FNTs, e que as FNTs constituem uma categoria fonética e acusticamente heterogênea. Nós chegamos a mencionar, no capítulo Revisão da Literatura, que as FNTs talvez servissem como uma sinalização para o ouvinte, dizendo algo como "preste atenção no que eu disse até agora, mas fique comigo porque ainda não terminei". Ou seja, as FNTs teriam uma função dupla e, aparentemente, contraditória: encerrar e continuar. As FTs, por outro lado, teriam como função primordial simplesmente encerrar, delimitar um enunciado, para que o ouvinte pudesse finalizar o processamento da ideia contida naquele enunciado e se preparar (ou não) para iniciar o processamento do próximo.

Pode-se conjecturar que as FTs não seriam propriamente inseridas, mas que seriam apenas repouso, passividade, falta de energia (pausa, ponto final na declinação da frequência fundamental). Ou seja, as FTs se caracterizariam por ausência, não por inserção de fenômenos prosódicos. As FNTs, por outro lado, seriam fenômenos prosódicos ativos, com uma dupla função de demarcação e continuidade; e seriam, efetivamente, inseridas. Nessa ótica, o fenômeno da deleção de fronteira prosódica não constituiria, propriamente, uma ausência de fenômenos prosódicos, mas uma substituição de uma determinada fronteira por manifestações acústico-fonéticas indicativas de continuidade discursiva, de que haveria adição de (pelo menos) uma nova palavra à(s) palavra(s) anterior(es) do período. Nessa perspectiva, não existiria deleção de fronteira, mas, sim, substituição de fenômenos prosódicos, uma inserção de características acústico-fonéticas que seriam perceptualmente interpretadas como continuidade do fluxo normal da fala.

5.5.2 Fala espontânea vs. “fala escrita”

Independentemente dessa discussão sobre a natureza de cada incongruência de fronteira, o que nos chamou atenção, como já dissemos, foi onipresença do fenômeno de inserção inapropriada de fronteiras prosódicas, seja para indicar "módulo de informação", seja para indicar a conclusão de um enunciado. Cada uma dessas inserções indicou uma dissonância entre o período escrito e a interpretação que lhe foi dada, em tempo real, pelo leitor, em prejuízo deste. Essa interpretação equivocada, que se mostrou bastante frequente, parece revelar uma tendência cognitiva no sentido de empacotar material linguístico, de agregar palavras em unidades, em constituintes, em enunciados. Na fala espontânea, onde quem define cada unidade de informação é o próprio falante, esse tipo de tendência não se manifesta; mas, no caso do texto escrito, temos dois problemas (pelo menos). O primeiro é que pode acontecer de o texto não ter sido escrito pelo leitor; e, mesmo que tenha sido, o texto foi escrito justamente porque o leitor não consegue guardar o material inteiro na memória; ou

seja, o leitor não tem o texto completo na cabeça. O segundo problema é que a modalidade escrita, além de ser heteromórfica com relação à modalidade falada, permite, em tese, amplitudes e profundidades linguísticas que não costumam se manifestar na fala espontânea

O que queremos dizer é que a escrita é produzida em uma escala de tempo e de permanência diferente da fala espontânea. A escrita, durante sua produção, pode ser relida, e burilada, e estendida, e pode adquirir camadas de complexidade que a fala espontânea não consegue e não costuma atingir. E, por isso, um texto escrito, mesmo quando é escrito para ser lido, costuma revelar a sua origem quando se transforma em fala lida. Um exemplo extremo seriam as *center-embedded clauses*, também chamadas de orações relativas de encaixe central recursivo. Períodos desse tipo não são produzidos espontaneamente, mas podem ser produzidos, a título de curiosidade, na escrita e, conseqüentemente, na fala lida. Uma fábula popular musicada por João de Barro conta a história de uma formiga cujo pé ficou preso na neve. A formiga "reza baixinho", enfileirando uma série de orações relativas encaixadas à direita, que, apesar de curiosas, não apresentam maior dificuldade de compreensão, pelo menos no nível sintático:

- Meu Deus, tu, que és tão forte que governas a morte que mata o homem que bate no cão que persegue o gato que come o rato que rói o muro que tapa o sol que derrete a neve, desprende meu pezinho!

No entanto, se reescrevemos o período com "orações relativas de encaixe central recursivo", obteremos um enunciado que é muito difícil (impossível?) de processar sintática e semanticamente, e de reproduzir por meio da fala espontânea:

- Meu Deus, tu, que és tão forte que a morte que o homem que no cão que o gato que o rato que o muro que o sol que a neve derrete tapa rói come persegue bate mata governas, desprende meu pezinho!

Reconhecemos que tal argumentação estende-se para um pouco além do razoável, mas o objetivo é demonstrar que a escrita permite construir constituintes e estruturas sintáticas e semânticas que excedem os limites do que seria considerado cognitivamente aceitável, ou cognitivamente usual, na fala espontânea (e, por extensão, na fala lida). Como mencionamos no capítulo Revisão da Literatura, citando uma analogia de Lobato (1986), certas operações aritméticas podem ser feitas facilmente no papel, mas muito dificilmente na cabeça.

É verdade que, via de regra, não percebemos as dificuldades que enfrentamos com textos escritos, até porque a maioria das pessoas não tem necessidade de lê-los em voz alta. Estudos têm mostrado, entretanto, que, mesmo na leitura silenciosa, estamos constantemente relendo trechos já lidos para consolidar algum aspecto de nossa compreensão. Beymer, Russell e Orton (2005) compararam a leitura silenciosa de textos em parágrafos largos (como este) com a leitura de parágrafos mais estreitos (como colunas de um jornal) e descobriram que há 0,56 releituras por segundo nos parágrafos largos, contra 0,36 em parágrafos estreitos. E analisando um banco de dados de rastreamento ocular de leitura silenciosa (COP; DIRIX; DRIEGHE; DUYCK, 2016), considerando apenas o arquivo "*Nineteen unbalanced Dutch (L1)–English (L2) bilingual*", nós contabilizamos a distribuição do número de fixações oculares por palavra durante a leitura e o ilustramos na **Tabela 63**.

Tabela 63 - Número de fixações por palavra

Número de fixações	Ocorrências
0	213400
1	287675
2	41838
3	5377
4	821
5	145
6	25
7	6
8	1
9	1
10	1
11	1

É possível notar que há um número muito grande de fixações adicionais em palavras já lidas. É fato conhecido, segundo Staub e Rayner (2007, p. 3), que "*in skilled readers, about ninety percent of saccades move the eyes forward, with the rest moving the eyes backward in text, either to resolve comprehension difficulty or to correct error in the programming of forward saccades*"¹. O leitor, então, efetua regressões oculares frequentes durante a leitura silenciosa; ele lê uma palavra, segue adiante e, muitas vezes, retorna e relê a mesma palavra. Staub e Rayner (2007), como mostramos, acreditam que muitos desses eventos de regressão são motivados por dificuldade de compreensão do que foi lido.

¹ Em leitores experientes, cerca de 90% dos movimentos oculares movem o olhar adiante no texto; os restantes 10% movem o olhar para trás, regressando a trechos anteriores do texto para sanar dificuldades de compreensão, ou corrigir erros de programação dos movimentos oculares. (tradução nossa)

Entretanto, não sabemos exatamente que tipos específicos de dificuldade foram esses. Será que o leitor leu uma palavra errado? Será que ele não compreendeu o sentido global de uma oração, ou de um enunciado? Ou será que ele não compreendeu um determinado trecho por não ter segmentado esse trecho corretamente e, por conseguinte, não ter sido capaz de reconstruir suas unidades discursivas, sejam módulos de informação, seja o enunciado completo?

Desconheço se essas perguntas já foram respondidas, mas nosso estudo pode ser analisado, também, sob essa perspectiva. Na fala lida, considerando o nosso corpus, a releitura em voz alta ocorreu apenas quando houve retratação, principalmente nos casos já discutidos de desvio de ortoépia e troca de palavra. Não acreditamos, entretanto, que, na leitura silenciosa, as regressões aconteçam apenas nesses casos. Como discutimos, as incongruências de fronteira indicam uma reconstrução defeituosa da segmentação do período e isso deve influenciar a compreensão do que foi lido. Como não observamos retratações de incongruências de fronteiras na fala lida, sugerimos, anteriormente, que a compreensão seria uma tarefa secundária para o leitor, e/ou que ele simplesmente optava por não se retratar. Considerando os dados de regressão em leitura silenciosa, não é impossível que parte dessas regressões possa relacionar-se a reconstruções de uma segmentação inicial defeituosa do período, e que essas regressões não se materializem como retratações na fala lida porque o leitor está mais preocupado com seu desempenho oral do que com a compreensão perfeita do que está lendo.

A essas conjecturas, soma-se outra. Se, com relação à fala espontânea, quando estamos na condição de ouvintes, não temos a possibilidade de reouvir (e nem reivindicamos, via de regra, essa possibilidade), por que, quando estamos na condição de leitores, relemos com tanta frequência? Parte da resposta deve estar relacionada à heteromorfia entre fala e escrita. E note-se que consideramos ‘heteromorfia’ não apenas a ausência da representação de determinados elementos da fala espontânea pela modalidade escrita, mas também as possibilidades construtivas que a escrita permite, mas que a fala espontânea limita. Como discutimos brevemente acima: a fala espontânea não permite o que a escrita permite e, assim, enquanto a escrita frequentemente se torna mais complexa do que a fala espontânea, a fala espontânea (não ensaiada), produzida em tempo real, está sempre sob controle das limitações cognitivas do falante, que são análogas, em média, às limitações cognitivas do ouvinte.

Se reunirmos as duas ideias acima, ficaremos com duas hipóteses. A fala espontânea é um produto completo, com elementos segmentares e suprasegmentares coerentemente organizados, 100% pronta para consumo; e, além disso, seu produtor/fornecedor (o falante) a submete a um controle de qualidade contínuo, em tempo real, substituindo imediatamente o produto quando julga que ele não

está adequado para consumo. A escrita, por outro lado, é um produto incompleto, que vem desmontado, faltando peças (componentes suprasegmentares) e sem um manual de instruções 100% amigável; e, além disso, como não é produzida em tempo real, pode oferecer dificuldades a uma reconstrução que deverá ser realizada, via de regra, em tempo real; seja na leitura silenciosa, seja na leitura em voz alta.

5.5.3 Produção & compreensão

Todavia, será que o falante, o produtor da fala espontânea, entrega um produto de tão fácil consumo apenas porque está empenhado em facilitar a sua utilização por terceiros (o ouvinte)? Ou não seria porque ele também é consumidor de seu próprio produto? Essa possibilidade já foi brevemente discutida neste trabalho, quando citamos Hartsuiker (2013). Esse autor critica a ideia de haver vias internas de monitoramento da fala e acredita que o monitoramento é feito a partir da fala externalizada. Ou seja, o falante monitora o mesmo produto que entrega ao ouvinte. E não apenas monitora, mas consome, toma conhecimento do que falou apenas depois de dizer. É uma posição análoga à defendida por Christiansen e Chater (2015, p. 6), quando dizem que "*we start by noting the powerful intuition that we speak 'into the void' – that is, that we plan only a short distance ahead*"¹. Posição, entretanto, que é criticada por Chacón, Momma e Phillips (2016), que sugerem que ela seria apenas uma versão da chamada "incrementalidade radical", que defenderia um planejamento da produção linguística palavra por palavra, em íntima e imediata conexão com sua articulação fonética; um modelo que, segundo eles, não se sustentaria empiricamente.

No entanto, quando nos alinhamos a Hartsuiker (2013), não estamos defendendo uma incrementalidade radical na produção, mas sugerindo uma incrementalidade radical na compreensão. Independentemente de como se dê o planejamento da fala espontânea, de que tipos ou níveis de representação cognitiva ou de vias de monitoramento estejam envolvidos nesse planejamento, o que estamos sugerindo é que todos os tipos e níveis de representação cognitiva da linguagem só são realmente compreendidos, percebidos concretamente, quando se transformam em fala espontânea, seja por meio de sua externalização sonora articulada, seja por meio da *inner speech*, nossa fala interna.

A própria existência da "fala interna" aponta nessa direção. Por que a fala interna seria tão onipresente, tão universal, como mostram Alderson-Day e Fernyhough (2015), se temos *loops* de

¹ Começamos ressaltando a poderosa intuição de que falamos “no vazio” – ou seja, a distância temporal entre o planejamento e a fala é muito curta. (tradução nossa)

monitoramento interno pré-articulatórios (considerando que a fala interna é articulada, mesmo que interna e silenciosamente)? Se realmente dispuséssemos de monitoramento semântico e sintático pré-articulatório, a produção de uma fala interna seria um exemplo de redundância cognitiva de difícil justificativa e, além disso, de surpreendente universalidade. Estendendo as ideias de Hartsuiker (2013), entendemos que não dispomos de tal monitoramento ou, pelo menos, que esse monitoramento seria bastante abstrato, não completamente interpretável e, assim, a fala (interna ou externa) seria o único produto concreto e 100% interpretável a nos revelar a maior parte de nossos "pensamentos".

Como disse Alexander von Humboldt, em seu *‘Cosmos: A Sketch of the Physical Description of the Universe’* (1845-1862): *But thought and language have ever been most intimately allied. If language, by its originality of structure and its native richness, can, in its delineations, interpret thought with grace and clearness, and if, by its happy flexibility, it can paint with vivid truthfulness the objects of the external world, it reacts at the same time upon thought, and animates it, as it were, with the breath of life.*¹ A fala produzida, portanto, seria a intérprete do pensamento, a tradutora, em termos palpáveis e perfeitamente compreensíveis (e necessários), de nossos níveis de representação cognitivos pré-linguísticos.

5.5.4 A incontornável memória de trabalho (WM)

Retomando o que discutimos até aqui, levantamos as seguintes hipóteses:

- o leitor não conhece o enunciado como um todo (conhece, talvez, às vezes, um par de palavras adiante do que está lendo em voz alta) porque tem limitação cognitiva para guardar, na memória, uma sequência de palavras com mais do que alguns poucos itens
- o leitor insere fronteiras incongruentes porque não conhece o enunciado como um todo — não tem ideia de sua extensão, nem de suas unidades de informação — e existe uma pressão cognitiva para empacotar constituintes, unidades de informação e enunciados

¹ Cosmos: Um rascunho da descrição física do Universo (1845-1862): Pensamento e linguagem sempre foram intimamente interligados. Se a linguagem, por sua originalidade de estrutura e sua inerente riqueza, pode, por seus meios de expressão, interpretar o pensamento com graça e clareza, e se, por sua exuberante flexibilidade, pode pintar com riqueza de detalhes os objetos do mundo externo; ela, a linguagem, age, ao mesmo tempo, sobre o próprio pensamento, e o anima, por assim dizer, com o sopro da vida. (tradução nossa)

- as dificuldades cognitivas do leitor — tanto a limitação que o impede de lidar com um número maior de palavras ao mesmo tempo, quanto a pressão cognitiva para delimitar constituintes, unidades e enunciado — derivam das características e limitações da memória de trabalho

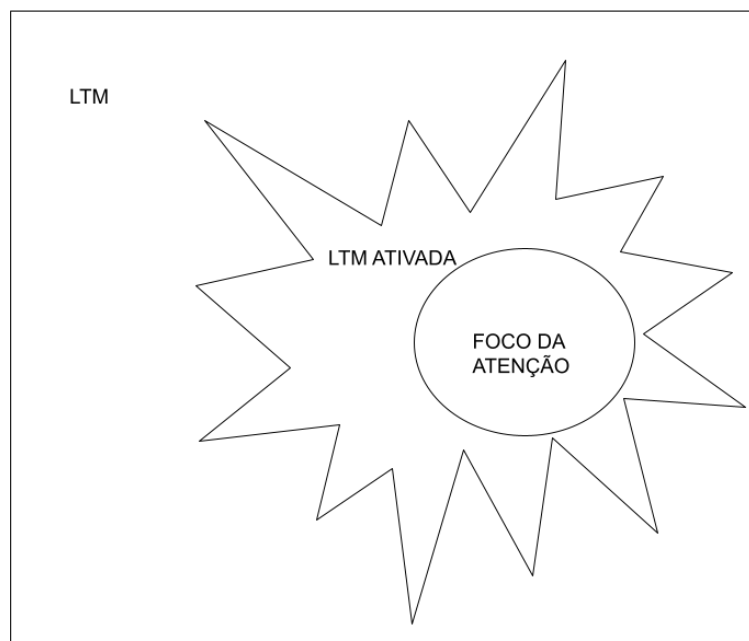
Uma hipótese mais geral para explicar as incongruências de fronteira prosódica é que a memória de trabalho impõe limitações que precisam ser superadas para que o desempenho linguístico seja apropriado. A pressão para empacotar unidades seria uma forma de superar essas limitações, pois unidades empacotadas, aglutinadas, que sofreram *chunking* (como discutimos no capítulo Revisão de Literatura), deixariam de ser uma série de palavras isoladas para se transformar em uma unidade semântica, um elemento único, liberando espaço na memória de trabalho (WM). Na sequência, aprofundaremos e alargaremos essa hipótese para incluir outros fenômenos linguísticos e cognitivos.

A tópicos memória de trabalho tem sido extensivamente estudado e a literatura sobre o assunto é bastante rica, propondo modelos e teorias para todos os gostos. Nosso objetivo não é criar um novo modelo, apesar de Watkins (1984, apud LOGIE; BELLETIER; DOHERTY, 2021) ter comparado teorias cognitivas a escovas de dente: todos precisam de uma, mas ninguém gosta de usar a do outro. Vamos, então, apenas recolher o melhor (em nossa visão) de algumas teorias para prosseguir em nossa discussão. O termo "memória de trabalho" parece ter sido criado por George A. Miller, na década de 60 (LOGIE; BELLETIER; DOHERTY, 2021), mas, antes disso, o conceito já existia. John Locke, em seu "*Essay Concerning Human Understanding*", do século XVII, já defendia que a retenção de conhecimento seria feita de duas maneiras: "*this is done in two ways. In the first, the idea is kept actually in view for some time — this is called ‘contemplation’*. *The second kind of retention is the power to revive again in our minds ideas that have come to us and then disappeared. This is **memory**, which is the store-house (so to speak) of our ideas*" ¹ (grifos nossos). Locke, então, chamava de "contemplação" o que hoje nos acostumamos a chamar de memória de trabalho. Tal abordagem, que não considerava a WM como memória, propriamente, mas como "contemplação", ou uma espécie de atenção, não é apenas uma curiosidade histórica. Hoje em dia, há teorias que modelam a WM, simplesmente, como uma ativação de setores da memória de longo prazo (LTM). A **Figura 32** ilustra

¹ Ensaio sobre a compreensão humana: Isto [a retenção] se faz de duas maneiras. Na primeira delas, a ideia permanece visível por algum tempo – isto se chama contemplação. Na segunda, recorre-se ao poder de reviver em nossas mentes, novamente, ideias que já estiveram visíveis, mas, então, desapareceram. Esta segunda maneira é chamada memória, que é o armazém (por assim dizer) de nossas ideias. (tradução nossa)

uma dessas teorias, a *embedded-processes approach to WM*, defendida por Cowan, Morey e Naveh-Benjamin (2021).

Figura 32 - *Embedded-processes approach to WM*



Observa-se na **Figura 32** que a WM está representada pela ativação da LTM (a área “estrelada”), incluindo o foco imediato da atenção. Observa-se também que a atenção pode estar voltada, por alguns momentos, apenas para uma fração da LTM ativada. O modelo pressupõe uma integração da WM com a LTM, considerando que a primeira depende da segunda e funciona sempre integrada a ela. Essa pressuposição é importante porque evidencia uma dificuldade para separar fenômenos atribuíveis exclusivamente à WM de fenômenos que, de alguma forma, utilizam-se de recursos da LTM.

Tal dificuldade para identificar a extensão da utilização da LTM pela WM parece estar na origem de boa parte das controvérsias quanto à real capacidade de retenção da WM. A pedra fundamental dessa discussão deve ter sido a declaração de Miller (1956) de que "o número mágico 7, mais ou menos 2", seria o limite da WM. Desde então, outros limites para a WM foram propostos e, dentre esses, nos alinhamos com o "*magical number 4*" proposto por Cowan (2001). Na realidade, não existe um número mágico, mas, de acordo com uma série de estudos analisados por Cowan (2001),

parece haver uma capacidade média em torno de 3 a 5 unidades (*chunks*), com uma variabilidade individual de, no mínimo, 2 unidades a, no máximo, 6 unidades.

Para chegar a esses números, foi necessário controlar, da melhor forma possível, a possibilidade de que, por exemplo, 2 unidades quaisquer na WM pudessem ter sofrido aglutinação (*chunk*) e se transformado em apenas 1 unidade, e que essa aglutinação não tenha sido considerada pelo pesquisador. Ou seja, foi necessário garantir que a contagem de unidades na WM não tenha sido falseada por algum tipo de associação entre essas unidades, pois, nesse caso, quaisquer números de unidades, uma vez aglutinadas, deveriam ser consideradas, para efeito de aferição da capacidade da WM, como apenas 1 unidade. E esse objetivo, segundo o autor, só poderia ser atingido respeitando certas condições experimentais cujo objetivo seria, justamente, impedir que o indivíduo conseguisse aglutinar elementos (unidades) da WM.

A dificuldade em manter mais do que 3 unidades independentes na WM é considerada por alguns como a causa para a dificuldade em processar orações relativas de encaixe central recursivo, como exemplificamos acima, na fábula da formiga. Quando dizemos (1) "Meu Deus, tu, que és tão forte que a morte que o homem mata governas, desprende meu pezinho!", não é tão difícil compreender o período. Se acrescentarmos, entretanto, uma oração relativa encaixada a mais, a dificuldade será praticamente intransponível: (2) "Meu Deus, tu, que és tão forte que a morte que o homem que no cão bate mata governas, desprende meu pezinho!". Postulamos que essa dificuldade em (2) derive da impossibilidade de manter na WM, ao mesmo tempo, "Deus", "morte", "homem" e "cão" (4 unidades), sem poder aglutiná-las para reduzir a carga da WM. Essa aglutinação ocorreria imediatamente caso surgisse um verbo que pudesse constituir, com qualquer desses nomes, uma unidade semântica: "Deus governa", "a morte mata", "o homem bate".

Note-se que estamos fazendo um raciocínio processual, não estrutural. Já argumentamos, nesta discussão, que uma análise a posteriori — *post mortem*? — não refletiria os caminhos percorridos pelo processamento da língua, que é, fundamentalmente, serial, sucessiva, produzida e analisada sequencialmente, unidade por unidade (quaisquer que sejam essas unidades: sílabas fonológicas, palavras, sintagmas, frases prontas...). Uma representação estrutural esquemática de um período seria, então, apenas uma abstração, uma evocação de determinadas relações estabelecidas a partir de um conhecimento global do período. Esse tipo de representação não deve refletir, adequadamente, o processo de produção e interpretação real da língua. Como disse Rambow (2010), árvores de estrutura frasal (*phrase structure tree*, PST) ou de dependência (*dependency tree*, DT) representam, apenas,

teorias sintáticas; e divergem, e se transformam, de acordo com as teorias que representam. Compreendemos que certas teorias defendem que a produção, nas línguas, parte da competência sintática (de estruturas sintáticas universais) para o desempenho (tipologicamente variável) e, sob essa perspectiva, faria sentido basear toda análise linguística em árvores de estrutura sintática. Reservamos, contudo, o direito de testar outras formas de análise e de privilegiar, nesses testes, uma abordagem menos estrutural e mais processual.

Vamos propor, então, um método de análise sequencial, processual, dos períodos (1) e (2), privilegiando a restrição cognitiva imposta pela WM e considerando que tal restrição seria uma das mais relevantes no processamento da linguagem (e, por tabela, na cognição, de um modo geral). Esse método — que não tem por objetivo representar um modelo cognitivo, mas, apenas, servir como auxílio ao raciocínio sobre as hipóteses que estamos levantando — representará as 3 unidades independentes manipuláveis pela WM e sua ocupação sucessiva pelo material linguístico apresentado em um determinado período. Cada par de parêntesis representará uma unidade independente na WM — ()() — e será ocupada, sucessivamente, pelo material linguístico do período; e as setas (->) indicarão cada passo hipotético do processamento. No caso do exemplo (1), teríamos, numa análise grosseira, de baixa granularidade, o seguinte:

(3) ()()-> (Deus)()-> (Deus)(que)()-> (Deus)(que)(a morte)-> (Deus)(que a morte)()-> (Deus)(que a morte)(que)-> (Deus)(que a morte)(que o homem)-> (Deus)(que a morte)(que o homem mata)-> (Deus)(que a morte que o homem mata)(governa)-> (Deus que a morte que o homem mata governa).

Se analisarmos, da mesma forma, o exemplo (2), teremos:

4) ()()-> (Deus)()-> (Deus)(que)()-> (Deus)(que)(a morte)-> (Deus)(que a morte)()-> (Deus)(que a morte)(que)-> (Deus)(que a morte)(que o homem)-> (Deus)(que a morte)(que o homem)que o cão*...

O asterisco indica que faltou uma casinha para o cão na WM, pois Deus, a morte e o homem já estavam ocupando espaços e não surgiu nenhuma informação que permitisse aglutinar nenhum dos elementos ocupantes da WM e reduzir sua ocupação. Numa referência infame, a aglutinação de duas unidades em apenas uma representaria uma situação análoga ao ditado "quem casa quer casa": quando dois elementos são aglutinados, eles "se mudam" para a mesma casa, o mesmo espaço, a mesma unidade na WM.

Não esqueçamos que não sabemos exatamente a extensão do que está na LTM, e que são informações armazenadas na LTM que permitem a aglutinação de elementos na WM. Pensamos que uma série de relações pré-fabricadas entre elementos, inclusive relações sintáticas esquemáticas, permitiriam aglutinações rápidas que reduziriam, imediatamente, a sobrecarga da WM. É possível, inclusive, pensar no processo descrito acima considerando um limite hipotético de apenas 2 unidades na WM, já que não podemos descartar que essa seja a real capacidade da WM quando nos referimos, especificamente, aos limites impostos à linguagem. Como vimos, o limite inferior de capacidade da WM, em indivíduos normais, é de apenas 2 unidades independentes e isso significa que — numa analogia computacional que seria melhor evitar — o "software língua" tem de ser projetado para rodar no "hardware cérebro" e esse hardware tem uma séria limitação de apenas 2 unidades independentes manipuláveis pela atenção ou, em última análise, pela consciência. No exemplo abaixo, baseado no período (1), exemplificamos o processo de ocupação da WM e de aglutinação de unidades considerando um limite de apenas 2 unidades independentes na WM:

(5) ()-> (Deus)()-> (Deus)(que)-> (Deus)(que a morte)-> (Deus que a morte)()-> (Deus que a morte)(que)-> (Deus que a morte)(que o homem)-> (Deus que a morte)(que o homem mata)-> (Deus que a morte que o homem mata)()-> (Deus que a morte que o homem mata)(governa)-> (Deus que a morte que o homem mata governa)().

Nesse exemplo, consideramos que a sequência "que o homem mata" é aglutinada imediatamente à unidade "morte", pois a LTM já teria essa estrutura sintática disponível para ir agregando elementos. Como dissemos, a extensão das associações pré-formadas disponíveis na LTM é desconhecida (e deve variar entre indivíduos), mas a questão fundamental é reconhecer que a WM impõe um limite à quantidade de unidades independentes (2, 3...) que podem ser manipuladas, e que as línguas são construídas apesar dessa limitação, em torno dessa limitação, e sem poder modificar, diretamente, essa limitação.

Um exemplo da variabilidade na extensão das associações pré-formadas na LTM e, conseqüentemente, na velocidade de processamento e aglutinação de unidades da WM, pode ser encontrado num trabalho de Leite (2012), que analisou a compreensão de textos lidos em voz alta por falantes de PB e português europeu. Entre seus achados, a autora identificou que "quanto maior a escolaridade do sujeito, menos ele precisa gastar recurso no nível de palavra para que haja a compreensão do sentido global do texto" (LEITE, 2012, p. 174). Essa conclusão foi obtida a partir da capacidade dos participantes de recordar a ocorrência de palavras específicas no texto lido, e essa

capacidade, surpreendentemente, foi maior nos sujeitos menos escolarizados. É verdade que a autora não deu aos seus achados a interpretação que estamos dando, mas acreditamos que seus resultados poderiam ser explicados dessa forma: sujeitos mais escolarizados "gastam menos recursos no nível da palavra" porque, nesses sujeitos, as palavras permanecem menos tempo na WM, sendo rapidamente aglutinadas em unidades maiores, compostas por mais de uma palavra, e assumindo imediatamente as características representativas de uma unidade semântica, descolada do nível fonológico de seus constituintes lexicais. Trata-se, em alguma extensão, de uma expertise adquirida com a prática da língua, automatizando determinados processos reiteradamente repetidos, que passam a estar inscritos na LTM e prontos para auxiliar no processamento da linguagem. E, como são processos mais automatizados, não prestamos tanta atenção a cada detalhe do processo e, no caso dos sujeitos mais escolarizados do estudo citado, os detalhes são, justamente, as unidades lexicais.

Considerar a WM como um fator limitante para a linguagem não é, exatamente, uma novidade. Mencionamos no capítulo Revisão da Literatura o modelo de processamento da linguagem proposto por Christiansen e Chater (2015) que considera o "*now or never bottleneck*" como um fator limitante fundamental da linguagem. Em resumo, os autores descrevem a compreensão da linguagem como uma corrida desesperada para aglutinar palavras e liberar espaço na WM (*chunk and pass*). Eles acreditam que esse é o gargalo mais importante no processamento da linguagem e que, para superá-lo, utilizamos três estratégias: processamento rápido, ansioso, "desesperado" (*eager processing*); vários níveis de representação (fonético, fonológico, lexical etc.); e ferramentas de predição. O modelo é interessante por ressaltar a importância da WM no processamento da linguagem, mas, em nosso entender, ele falha em, pelo menos, alguns pontos, quais sejam:

- Os autores acreditam que uma informação que é recodificada para um nível superior de representação deixa de interferir com as informações que permanecem no nível inferior de representação. Por exemplo, eles acreditam que sílabas fonológicas, uma vez aglutinadas e representadas como palavras, não interferem mais com as demais sílabas fonológicas, uma vez que deixaram de ser sílabas isoladas e se tornaram uma palavra. Nosso entendimento é que qualquer unidade independente presente na WM interfere, compete por espaço, com outras unidades.
- Eles dizem que a produção segue o caminho inverso da compreensão, partindo de unidades (*chunks*) de nível discursivo e chegando até o nível fonético e articulatório, e acrescentam que, para liberar a WM o mais rapidamente possível, o processo de "fragmentação" dos *chunks* discursivos (para chegar, ao final do processo, à articulação fonética) ocorreria tão logo esses

chunks discursivos estivessem disponíveis na memória. Eles usam esse raciocínio para explicar dados da literatura que apontariam que objetos sintáticos "mais leves" seriam produzidos, preferencialmente, antes de objetos "mais pesados" (mais complexos, maiores). Entretanto, Yamashita e Chang (2001) já demonstraram que esse padrão de *light objects first* não se aplica, ao japonês, uma língua SOV. Em japonês, objetos longos tendem a ser posicionados antes de objetos mais curtos, comprometendo a universalidade da proposta de Christiansen e Chater (2015).

- Outra proposta dos autores é a de que as informações linguísticas sofrem, na compreensão, o que eles chamam de "*lossy compression*", uma compactação com perda de informações. Nós achamos que, fundamentalmente, não é isso que acontece. Haveria *lossy compression* se o objetivo do ouvinte/leitor fosse recuperar cada vírgula do que foi dito/escrito, mas acreditamos que não é esse o objetivo da linguagem. O objetivo da linguagem é, teleologicamente, transmitir uma mensagem (para manter um foco estreito, ignoramos, aqui, outras possíveis funções da linguagem). Pode-se dizer que, onde os autores enxergam *lossy compression*, nós vemos um processo de semântica composicional e, de acordo com o princípio do contexto, de Frege, "*it is enough if the sentence as whole has meaning; thereby also its parts obtain their meanings*"¹.

Além disso, os autores enfatizam as dificuldades relacionadas, principalmente, à compreensão da linguagem, e toda a sua proposta é construída a partir desse ponto de partida. Acreditamos que, talvez, a produção exija mais do sistema cognitivo do que a compreensão. Apesar de considerarmos produção e compreensão como fenômenos interligados, a produção impõe à WM as mesmas dificuldades que a compreensão, e ainda enfrenta dificuldades adicionais. Um exemplo prosaico de que a compreensão não costuma ser tão limitante pode ser observada nos tempos atuais. Com a pandemia do SARS-CoV-2, as pessoas têm ficado mais tempo em casa e têm recorrido mais a atividades via computador e internet. Crianças e jovens têm assistido aulas on-line, e é muito comum que assistam essas aulas em velocidade aumentada. Isso tem acontecido no mundo todo e em várias situações, e já existe até um termo para designar o fenômeno: *speed watching*. De fato, é uma experiência que todos podemos fazer, aumentando a velocidade de áudios e vídeos em até duas vezes,

¹ É suficiente que o período como um todo tenha sentido; assim, também suas partes obtêm sentido. (tradução nossa)

ou mais, acima da velocidade real e, até certo limite, sem ter nenhuma dificuldade de compreensão; e, talvez, em algumas situações, até mesmo facilitando essa compreensão.

Maryellen C. MacDonald (MACDONALD, 2013) também defende que os limites cognitivos de processamento da língua afetam, sobretudo, a produção. Ela acredita que são as escolhas feitas pelos falantes durante a produção da língua que moldam suas características. Ou seja, a engenharia inconsciente e ininterrupta da língua operaria para facilitar a produção, e não a compreensão. A maioria das características observadas nas diversas línguas, então, teria derivado de adaptações para atingir maior ergonomia cognitiva para o falante, não para o ouvinte.

A "ergonomia cognitiva" mantém a língua operando dentro de certos limites, impostos, entre outros pela WM. Nosso ponto de vista é de que a WM impõe uma restrição fundamental, talvez a mais fundamental de todo o sistema cognitivo humano. De fato, é gigantesca a dissonância entre uma WM tão limitada e uma LTM, aparentemente, ilimitada. É absolutamente surpreendente que, com apenas 3,5 unidades independentes manipuláveis pela atenção, pela consciência imediata, o ser humano consiga fazer tudo o que faz, cognitivamente falando. Conseguir explicar as estratégias utilizadas pela cognição, em geral, e pelos sistemas linguísticos, em particular, para superar esse fosso quantitativo que existe entre WM e LTM parece-nos uma questão central.

Pode-se elaborar um raciocínio que consideraria, para o desenvolvimento diacrônico das estruturas linguísticas de uma língua, a existência de três tipos de caminhos a seguir: um fácil, um difícil e um impossível. Os caminhos impossíveis seriam impossíveis pelas restrições impostas pela cognição e, naturalmente, nunca foram percorridos. Os caminhos fáceis seriam aqueles que todas as línguas percorreram pois, para seus falantes, a utilização cotidiana da língua materna não impõe qualquer dificuldade intrínseca; e, na fase de aquisição da língua pelas crianças, os mesmos caminhos fáceis foram percorridos, pois a aquisição costuma ser feita sem esforço consciente, sem traumas e desconfortos, e dentro dos limites cognitivos e dos parâmetros da língua nativa. Resta-nos elaborar um pouco mais sobre os caminhos difíceis.

Quando discutimos as "orações relativas de encaixe central recursivo" (*center-embedded*) (CE), sugerimos que poderia ser impossível compreender períodos com um número de argumentos independentes não aglutináveis que ultrapassassem a capacidade da WM. E utilizamos um método arbitrário para tentar ilustrar a ocupação da WM pelos elementos do período. Apesar de as CEs serem, na literatura, o paradigma da influência da WM sobre as possibilidades de uma língua, podemos,

tentativamente, sugerir outros exemplos de construções linguísticas que também "forçariam a amizade" com a WM.

O primeiro exemplo será nominal e o segundo, verbal; e ambos empregam estruturas recursivas, como as CEs. E, nesse ponto, vale a pena lembrar o que se entende por recursão. Corballis (2014), em seu livro *The Recursive Mind: the origins of human language, thought, and civilization*, lembra que Steven Pinker e Ray Jackendoff definem recursividade (recursão) como um procedimento que convida a si mesmo, ou um constituinte que contém um constituinte do mesmo tipo. E Noam Chomsky (2014, p. 12), em *Minimal Recursion: Exploring the Prospects*, resolve reduzir a importância da recursividade e considerá-la simplesmente como *merge*: "*Gallistel [...] suggests that the “great leap forward” might have been externalization, not recursive generation (let us say just Merge)*"¹. *Merge*, em nossa concepção, é o mesmo que aglutinação (*chunking*) de duas unidades (itens lexicais ou aglutinações prévias de itens lexicais) em apenas uma. Por meio da simples aglutinação sucessiva, independentemente de haver recursão ou não, seria possível atingir qualquer objetivo semântico. E os princípios dessa composicionalidade seriam a produtividade e a sistematicidade, permitindo aglutinações de extensão e profundidade infinitas, respeitando o sistema de regras composicionais da língua (SZABÓ, 2020).

Retomando os exemplos de construções linguísticas que, apesar de respeitarem as regras composicionais da língua, não são produtivas – pois, supomos, esbarram nos limites da WM —, começamos com uma construção nominal recursiva:

(6) Porque és o avesso do avesso do avesso do avesso.

A composição semântica do que significa "o avesso do avesso do avesso do avesso", de fato, não é automática, pois, primeiro, é preciso fazer uma operação não linguística para reconhecer que "avesso do avesso" é, simplesmente, mais do mesmo, a mesma coisa, um retorno à condição original. Mas, mesmo reconhecendo isso, não é fácil estender o raciocínio para as reiterações de "avesso" e, assim, aglutinar as unidades lexicais em uma unidade semântica que represente o enunciado. Acreditamos que essa dificuldade provenha do fato de que as aglutinações, nesse caso, devem ser feitas de trás para a frente. Por exemplo, se numerarmos os "avessos" na ordem em que aparecem (1, 2, 3 e

¹ Gallistel [...] sugere que o “grande salto adiante” pode ter sido a externalização, e não a geração recursiva (chamemos, simplesmente, de ‘Merge’). (tradução nossa)

4), a aglutinação de 1 e 2 só pode ser feita depois da aglutinação de 2 com 3, que, por sua vez, só pode ser feita depois da aglutinação de 3 com 4. Dessa forma, é necessário manter 4 "avessos" na WM antes de iniciar, de trás pra frente, as aglutinações: (avesso 1)(avesso 2)(avesso 3) avesso 4*... Em um período com 3 "avessos", a operação seria muito mais natural: (avesso 1)(avesso 2)(avesso 3)-> (avesso 1)(avesso do avesso)()-> (avesso do avesso do avesso)()().

Trata-se de uma situação análoga a uma operação aritmética com prioridades de cálculo. Em $2+2+2+2$, aglutinamos cada novo "2" ao "2" anterior, não sendo necessário ocupar mais de duas unidades na WM: (2)(2)-> (4)()-> (4)(2)-> (6)()-> (6)(2)-> (8)(). Em $2+2*2^2$, no entanto, temos de fazer, primeiro, a última operação, 2^2 , depois $2*(2^2)$, e só então $2+(2*(2^2))$, e assim, teríamos, na WM: (2)(2)(2)2*...

Voltando para a língua, teríamos também, como exemplo de recursividade nominal em que a aglutinação dever ser feita de trás para frente, o exemplo de relações recursivas de parentesco/intimidade:

(7) O amigo do irmão do cunhado da minha namorada.

Essa frase equivale a:

(8) O cunhado da minha namorada tem um irmão cujo amigo...

O período (8) não apresenta nem uma fração da dificuldade de (7), o que sugere que a dificuldade não está associada ao número de unidades a serem aglutinadas — que, em ambas as situações, é 4 —, mas à necessidade de manter unidades não aglutinadas na WM. Em (8), o esquema seria: (cunhado)(da namorada)()-> (cunhado da namorada)(tem irmão)()-> (irmão do cunhado da namorada)(tem amigo)()-> (amigo do irmão do cunhado da namorada)()(). Em (7), por outro lado, o esquema seria (irmão)(do amigo)()-> (irmão)(do amigo)(do cunhado)-> (irmão)(do amigo)(do cunhado)da namorada*...

Outro exemplo que podemos analisar é verbal, como as CEs. Trata-se das construções causativas. Gomes (2012) fornece um exemplo de construção causativa em Mundurucu, uma língua indígena do tronco Tupi:

(9) Pedro akurice o' = jo - wuy

Pedro cachorro 3Sa = NCNT - lavar.PRF

‘Pedro lavou o cachorro’

(10) Maria Pedro be akurice o’= jo - wuy - at

Maria Pedro POSP cachorro 3Sa

‘Maria fez Pedro lavar o cachorro’

Em (10), a uma estrutura convencional do tipo SOV, acrescenta-se outro sujeito, que deu causa ao evento descrito pela estrutura SOV. De fato, não há dificuldade, em PB, para compreender "Maria fez Pedro lavar o cachorro". E, em Munduruku, também não deve haver nenhuma dificuldade. O esquema, em Munduruku, pode ser: (maria)()()-> (maria)(pedro be)()-> (maria)(pedro be)(akurice)-> (maria)(pedro be akurice)()-> (maria)(pedro be akurice)(o’= jo - wuy - at)-> (maria)(pedro be akurice o’= jo - wuy - at)()-> (maria pedro be akurice o’= jo - wuy - at)()().

Se acrescentarmos mais um elemento à construção, teremos um sujeito que dá causa ao sujeito que dá causa ao evento descrito pela estrutura SOV, e começamos a ter dificuldade. Em PB, teríamos Paulo fez Maria fazer Pedro lavar o cachorro". Em Munduruku, suponhamos, com o perdão dos estudiosos e dos falantes da língua, que a construção, caso exista, seja:

(11) Paulo Maria be Pedro be akurice o’= jo - wuy - at

O esquema para essa construção poderia ser:

(Paulo)()()-> (Paulo)(Maria be)()-> (Paulo)(Maria be)(Pedro be)-> (Paulo)(Maria be)(Pedro be)akurice*...

O que observamos é que, em causativas com mais sujeitos causadores, deve faltar espaço na WM para representar unidades independentes que deverão ser aglutinadas. Observemos que não importa muito se nossa construção em Munduruku está correta, tampouco se os esquemas propostos para a ocupação sucessiva de espaços na WM correspondem à realidade. O importante é que, a depender de como a língua organiza suas estruturas na LTM, a sucessão de elementos que vão ocupando a WM ao longo de um período (uma sucessão de palavras) será organizada e aglutinada de forma a não ultrapassar o limite da WM. Em línguas SOV, como o Munduruku, o Japonês e metade

das línguas do mundo, é bem possível que a LTM permita a aglutinação de S+O em uma unidade SO, pressupondo que os dois elementos estão unidos em um evento abstrato em que S age sobre O. Esse tipo de estrutura pré-fabricada na LTM permitiria que S e O não permanecessem separados, ocupando 2 unidades na WM, enquanto aguardassem o verbo (V), que os aglutinaria. O esquema padrão, então, seria: (S)()()-> (S)(O)()-> (SO)()()-> (SO)(V)()-> (SOV)()(). Em línguas SVO, por outro lado, se acreditarmos nas inúmeras teorias de que O é argumento interno do verbo, portanto mais intimamente ligado a ele, e que S é argumento externo, teríamos: (S)()()-> (S)(V)()-> (S)(V)(O)-> (S)(VO)()-> (SVO)()(). Não pretendemos argumentar contra tais teorias, mas talvez seja mais fácil concordar com um esquema em que S e V se aglutinassem antes da chegada de O: (S)()()-> (S)(V)()-> (SV)()()-> (SV)(O)()-> (SVO)()(). Do ponto de vista da WM, a LTM lhe prestaria melhores serviços se permitisse a aglutinação de S e V, em vez de ficar aguardando um possível O.

De qualquer forma, independentemente das características das estruturas pré-fabricadas pela LTM para facilitar o trabalho da WM, parece que continuam existindo 3 unidades independentes na WM que precisam ser aglutinadas, seja nas línguas SVO, seja nas SOV. Ueno e Polinsky (2009) compararam línguas SVO (inglês e espanhol) com línguas SOV (turco e japonês), analisando a incidências de dois vieses conhecidos — *pro-drop* (omissão do sujeito) e intransitividade (empregar verbos com valência 1, em desfavor de verbos com valência 2) – e descobriram que não há diferença entre as línguas SVO e SOV quanto ao viés *pro-drop*, mas que há um maior viés de intransitividade nas línguas SOV. A justificativa seria que, nessas línguas, haveria uma tendência a utilizar, no dia a dia, verbos intransitivos, para reduzir o número de argumentos pré-verbais e facilitar o processamento (evitar deixar duas unidades independentes ocupando espaço na WM, enquanto se aguarda o verbo).

Os exemplos iniciais que escolhemos evidenciam construções linguísticas que sobrecarregariam a WM. A partir de certo ponto, torna-se impossível processar os períodos da maneira habitual. A maneira habitual é, por assim dizer, indolor e recompensadora. Comparando a leitura de um período habitual da língua, como "o rato mordeu o gato que arranhou o cão que acabou fugindo" com uma construção do tipo "o cão que o gato que o rato mordeu arranhou acabou fugindo", percebe-se que, neste último, além de termos dificuldade para entender quem mordeu, quem arranhou e quem fugiu, falta (pelo menos para este autor) uma sensação de ter atingido a *gestalt*, de ter ficado suficientemente satisfeito com a composição semântica e sintática do período a ponto de poder seguir adiante e enfrentar novos períodos. No período anterior, habitual, mesmo que algum aspecto da composição semântica revele-se, depois, inapropriado (pergunta: "quem mordeu o gato?" resposta: "o cão?"), a "sensação" é de completude.

Entretanto, acreditamos que a língua poderia, se fosse pressionada para tanto, criar estruturas na LTM para processar algumas dessas construções “difíceis” com o objetivo de, pelo menos, tornar os casos limítrofes processáveis. Por exemplo, com treino e dedicação, um falante criaria estruturas na LTM que permitiriam que ele processasse com facilidade "o cão que o gato que o rato mordeu arranhou fugiu", mas não "o cão que o gato que o rato que a pulga picou mordeu arranhou fugiu". E, de qualquer forma, se a LTM criasse estruturas para processar a primeira sentença e não conseguisse criá-las para processar a segunda, é possível que tais sentenças se tornassem análogas a sentenças com sujeitos ou objetos múltiplos, como "o cão, o gato, o rato e a pulga morderam a criança, a mulher, o homem e o papagaio". Nesse tipo de período, o processamento — a apreensão da estrutura da sentença e do local ocupado por cada um de seus elementos — ocorre sem dificuldade; o difícil é lembrar, depois, cada um dos sujeitos e dos objetos.

Dados de um estudo de Morford (2003) parecem apontar nessa direção. Ele estudou indivíduos surdos que aprenderam sua primeira língua (língua de sinais) apenas na adolescência e que apresentavam dificuldades persistentes para interpretar e produzir estruturas sintáticas mais complexas. Esse é um fenômeno habitualmente observado em indivíduos que adquirem a linguagem tardiamente, tanto com relação a língua oral (crianças abandonadas ou maltratadas), quanto com relação à língua de sinais. Entretanto, o que os experimentos feitos por Morford mostraram foi que os déficits desses indivíduos não eram causados por dificuldades com o aprendizado da gramática das estruturas sintáticas mais complexas, mas com o processamento dessas estruturas. Quando ele reduziu as demandas de processamento, os indivíduos conseguiram melhorar bastante seu desempenho. Isso parece indicar, como mencionamos, que a construção de estruturas abstratas na LTM facilita o processamento, reduzindo a pressão sobre a WM e sobre a atenção, e acelerando a aglomeração de unidades linguísticas. E, nos sujeitos dos estudos, as estruturas abstratas não estavam suficientemente estabelecidas na LTM para facilitar o processamento da língua, apesar de eles conhecerem as regras gramaticais aplicáveis em cada situação. É possível que determinadas estruturas só possam ser estabelecidas na LTM durante a fase de aquisição da linguagem, em idades precoces, em fase anterior à poda sináptica.

E por que as línguas não são pressionadas na direção de conseguir trabalhar com um número maior de elementos? Talvez porque não haja motivação para isso. Parece que a composicionalidade das línguas se restringe, na prática tipológica das línguas conhecidas, a duas possibilidades. A primeira seriam as aglutinações por adjunção simples. Nesta situação, a unidade semântica "cão feroz" seria formada por simples adjunção de "feroz" a "cão": (cão)()()-> (cão)(feroz)()()-> (cão feroz)()(). Uma

unidade desse tipo daria origem a uma unidade de informação, na nomenclatura proposta por Izre'el (2018). A segunda possibilidade seriam as resoluções de estruturas argumentais de verbos, em que a formação de uma unidade semântica dependeria, via de regra, da aglutinação de 3 elementos: S, V e O. Essa unidade daria origem a um enunciado, na nomenclatura proposta pelo mesmo autor.

Com a utilização dessas duas estratégias — com limites hipotéticos de 2 e 3 unidades na WM, respectivamente, para casos de adjunção e de resolução de estrutura argumental — estariam satisfeitas todas as necessidades de composicionalidade semântica das línguas. Em condições naturais, não haveria impulso para criar construções linguísticas mais complexas, pois a manipulação cognitiva dessas estruturas mais complexas exigiria um esforço a mais para elaborar estruturas abstratas na LTM. Apenas por meio dessas estruturas abstratas na LTM, a WM conseguiria lidar com uma estrutura argumental de, digamos 4 ou 5 elementos, da mesma forma como lida com palavras de 4, 5 ou mais sílabas. Ao que tudo indica, a capacidade da WM é inelástica, enquanto a da LTM, não é.

O que propomos, então, é que as inserções de fronteira são tão frequentes porque existe uma pressão cognitiva para encerrar constituintes e aglutiná-los, com o objetivo de liberar espaço na WM. Essa aglutinação pode ser de unidades de informação, dando origem a unidades semânticas parciais; ou pode ser uma aglutinação das subunidades de um enunciado, dando origem a uma unidade semântica completa, que descreve um evento com “autonomia pragmática” e “interpretabilidade em isolamento”. Quando os elementos do enunciado se aglutinam em uma unidade semântica completa, o evento descrito deixa a WM e é registrado na LTM, liberando completamente o foco de atenção para novos desafios impostos pela língua.

5.5.5 Hipóteses adicionais

Para finalizar, levantaremos três hipóteses. Elas decorrem da análise dos dados obtidos por este estudo, mas, por terem surgido tardiamente no transcurso do trabalho, não houve tempo hábil para elaborá-las melhor ou aprofundar, como julgamos que mereceriam, sua validação. De qualquer forma, não poderíamos finalizar esta dissertação sem, ao menos, mencioná-las.

5.5.5.1 Hipótese I

Dissemos anteriormente que o fator limitante no processamento da linguagem poderia ser a produção, não a compreensão. Dissemos também que estávamos alinhados aos que defendem um monitoramento externo da produção linguística; ou seja, o falante só poderia saber realmente o que

disse após tê-lo dito. Nessa ótica, o falante também exerceria um papel de ouvinte da própria fala, de consumidor de sua própria produção linguística.

Segundo Barbosa (2019), a prosódia não é *o que se diz*, mas *como se diz*, e uma de suas funções seria segmentar o fluxo da fala. Discutem-se na literatura os motivos que levariam à existência dessa segmentação. Não discutimos aqui as diversas propostas existentes (sintática, informacional, cognitiva etc.), mas propomos (em possível alinhamento parcial com algumas delas) que:

- a) A prosódia (de segmentação) é produzida, primeiramente, em benefício do falante, não do ouvinte;
- b) O falante segmenta sua produção linguística por dois motivos: b1) para compreender o que planejou dizer e saber o que realmente disse; e b2) para compreender o que está dizendo, à medida que vai dizendo;
- c) Em grande parte, a segmentação produzida pelo falante reflete as unidades cognitivas que ele está produzindo. Os locais onde o fluxo da fala é segmentado – considerando as fronteiras não terminais e terminais — devem indicar onde ocorreu um processo de aglutinação de palavras, que passaram a ocupar uma mesma unidade na WM;
- d) A segmentação prosódica do fluxo da fala é uma consequência direta da WM; é causada pela WM; existe por causa das limitações da WM; e sua ocorrência reflete fenômenos de aglutinação de unidades que, por sua vez, são reflexo de estruturas abstratas previamente registradas na LTM;
- e) O falante pode, secundariamente, utilizar-se de fenômenos de segmentação prosódica em benefício do ouvinte, mas, como regra, os ouvintes beneficiam-se da segmentação prosódica por conta própria; porque aprenderam (como ouvintes, mas, principalmente, como falantes) o que essa segmentação significa, do ponto de vista da língua; e
- f) Parece-nos que, onde houver (e apenas onde houver) incongruência entre segmentação prosódica e organização sintática, será importante lembrar que sintaxe é um construto teórico, uma abstração, enquanto a prosódia (independentemente do que signifique) é uma realidade linguística concreta, manifesta, evidente.

5.5.5.2 Hipótese II

Payne (1997) considera que existem 3 categorias de relações gramaticais: sujeito, objeto e objeto indireto; e que talvez existam mesmo apenas 2, se encararmos o objeto indireto como uma categoria gramatical periférica, oblíqua, não indispensável à estrutura argumental do núcleo verbal. Ele acredita que essa restrição talvez decorra de limitações cognitivas para contabilizar um número maior de participantes em eventos descritos por verbos.

O projeto ValPal — *Valency Patterns Leipzig* — (HARTMANN; HASPELMATH; TAYLOR, 2013) fez um levantamento de valências verbais em diversas línguas a partir de uma lista padrão de verbos semanticamente representativos. Uma análise rápida de seu banco de dados on-line revela que o número de “participantes” da estrutura argumental básica desses verbos variou de 0 a 4. Não houve nenhum registro de padrões gramaticalizados de estrutura argumental para 5 participantes. Além disso, de todos os esquemas diferentes de estruturas argumentais básicas identificados (excluídas as alternâncias verbais), apenas 3% correspondiam a esquemas para 4 participantes (ex.: *The man tied the horse with a rope to the tree*)¹.

O que se nota em valências maiores do que 2 é que passamos a encontrar elementos que ressuscitam a discussão sobre argumentos não obrigatórios *versus* adjuntos; e recaímos no problema que abordamos anteriormente, e que nos levou a separar os participantes dos eventos descritos pelos verbos em participantes obrigatórios e não obrigatórios, simplesmente. Como dissemos, acreditamos que participantes não obrigatórios não participam da predicação primária; não são imprescindíveis à formação de uma unidade semântica autônoma, que poderia ser interpretada, em situação pragmática adequada, como um enunciado. Ou seja, não são participantes, mas apenas coadjuvantes.

A unidade semântica de determinado material linguístico é pré-requisito para que esse material, por variado que seja, possa ser manipulado pela WM como um elemento único. Quando essa unidade semântica é completa, quando tem “autonomia pragmática” e “interpretabilidade em isolamento”, além de ser um único elemento, pode ser transferida para a LTM e liberar espaço para que novo material linguístico seja manipulado pela atenção. Essa unidade semântica completa pode ser representada por uma construção verbal, ou não. No primeiro caso, temos o verbo e seus participantes obrigatórios,

¹ O homem amarrou o cavalo a uma árvore com uma corda. (tradução nossa)

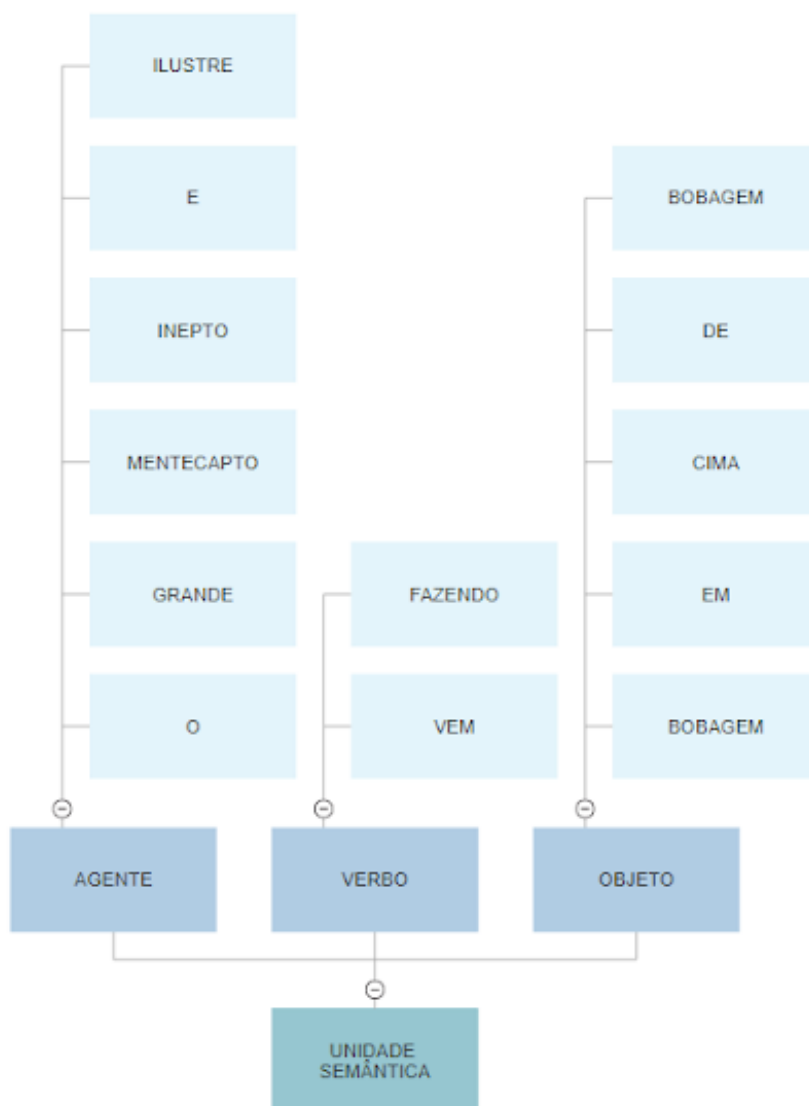
como analisado pelo ValPal. No segundo caso, temos qualquer palavra que, no contexto adequado, seja suficiente para possuir autonomia e interpretabilidade.

Wittgenstein (1953 apud IZRE'EL, 2018) analisa a diferença entre períodos completos (com verbos e seus participantes obrigatórios) e construções linguísticas curtas, “incompletas”, que seriam versões “encurtadas” das formas completas. Seriam frases como “horrível”, que no contexto apropriado poderiam corresponder a períodos como “essa casa é horrível”, “acho essa casa horrível”, “acredite em mim, conheço essa casa e ela é horrível”, “não adianta você dizer o contrário: essa casa é horrível” e assim por diante. Wittgenstein questiona o porquê de não chamar as formas completas de “versões alongadas da sentença”, em vez de chamar aquela (“horrível”) de “versão encurtada da sentença”. De fato, “horrível”, no contexto adequado, pode ser uma expressão perfeitamente autônoma e interpretável, correspondendo a uma unidade semântica completa.

Essa discussão nos leva ao ponto onde queremos chegar: a hipótese II. Essa hipótese é que nossa WM não nos permite lidar com construções linguísticas que ultrapassem o limite de 3 elementos independentes, a não ser que consigamos aglutiná-los em uma unidade semântica; e, em função disso, existiria um limite máximo para o número de participantes obrigatórios da estrutura argumental de qualquer verbo, que deve ser 2. A aglutinação de palavras (consideradas unidades semânticas, por si sós) costuma ocorrer por simples *merge*, composição semântica imediata de duas unidades sucessivas. A estrutura argumental de verbos, em que consideramos ser possível uma valência máxima de 2, implica que a unidade semântica só é atingida a partir da aglutinação de 3 elementos, que chamaremos de sujeito (ou agente, ou qualquer outro nome), verbo e objeto (ou paciente, ou o que mais ele puder se chamar). Nesse caso, haveria 3 unidades livres e independentes na WM aguardando aglutinação. Como discutimos acima, é possível que a LTM tenha estruturas pré-fabricadas que permitam unir duas dessas unidades enquanto aguarda-se a próxima (ex.: sujeito + verbo, aguardando o objeto), e nesse caso, estaríamos lidando, novamente, com um simples caso de *merge* de unidades adjacentes, não sendo necessário manter mais do que 2 unidades na WM. Entretanto, considerando que não costumamos ver verbos com valência maior do que 2 (considerando apenas os “argumentos obrigatórios”), é mais provável que os participantes do evento descrito pelo verbo (e o próprio verbo) permaneçam ocupando unidades na WM até que sejam aglutinados em uma unidade semântica completa.

A **Figura 33** ilustra esse raciocínio para o período “o grande mentecapto inepto e ilustre vem fazendo bobagem em cima de bobagem”, onde, no fim das contas, restam 3 unidades na WM (sujeito, verbo e objeto) para serem aglutinadas em uma unidade semântica completa.

Figura 33 – Aglutinação de palavras em unidades semânticas



A limitação de 3 unidades independentes é reflexo da limitação da WM. As línguas humanas foram (e estão sendo) construídas considerando essa limitação. Trata-se, na verdade, de uma limitação cognitiva básica, fundamental, que afeta toda e qualquer experiência humana. Na **Figura 34**, por exemplo, há mais de um evento ocorrendo ao mesmo tempo; há mais de um participante para cada

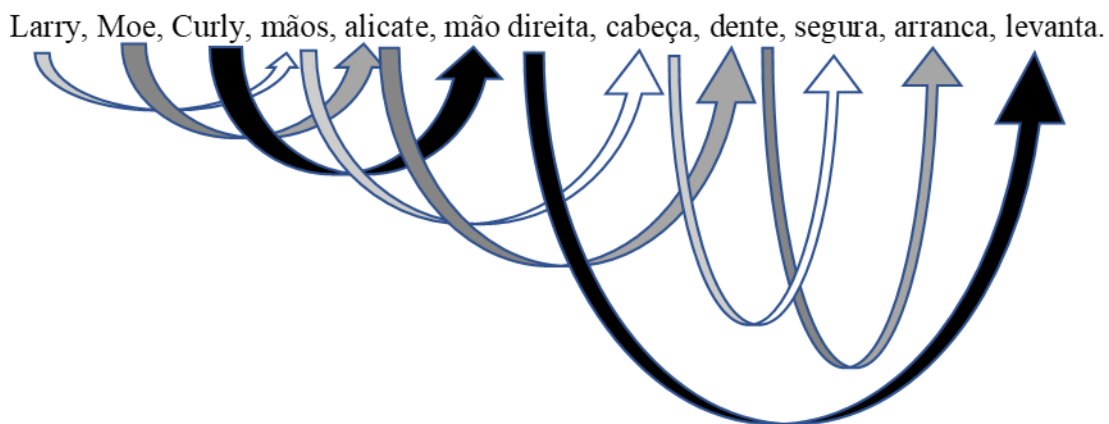
evento; trata-se, na verdade, de um evento complexo, e não temos recursos linguísticos para descrevê-lo de forma sintética, holística, capturando a totalidade do evento por meio de um gesto linguístico único. Da mesma forma que não conseguimos fixar a visão nos rostos dos três indivíduos ao mesmo tempo (só conseguimos fixa-la em um de cada vez), não conseguimos construir uma sentença que capture o evento como um todo. O que fazemos é capturar um elemento de cada vez e, sequencialmente, por meio de *merges* e estruturas argumentais de verbos que descrevem eventos com dois participantes, compor a cena, tijolo por tijolo: *Larry está segurando a cabeça de Curly, enquanto Moe usa um alicate para arrancar um dente de Curly, e Curly está fazendo uma careta e levantando a mão direita...* Ou seja, nós até compomos uma visão geral da cena em nossa LTM, mas nosso foco de atenção é limitado a um elemento de cada vez, e as línguas humanas acompanham tal limitação.

Figura 34 – Um evento complexo, com múltiplos agentes e ações simultâneas



Mesmo com foco de atenção (e, associadamente, de visão) tão limitado, se tivéssemos uma WM de maior capacidade, poderíamos construir estruturas linguísticas menos dependentes da focalização em pequenos trechos do evento. É possível que produzíssemos estruturas em que uma série de elementos não conectados fossem enfileirados e as aglutinações só acontecem tardiamente, com resultado final idêntico. Como, por exemplo, na **Figura 35**.

Figura 35 – Hipótese de possível período livre das limitações da WM



E por que a WM não evoluiu, não se expandiu, para facilitar o funcionamento da língua e de outras atividades cognitivas? Provavelmente, como apontam Lotem, Kolodny, Halpern, Onnis e Edelman (2016), porque não houve pressão evolutiva para tal. A língua e a cognição foram construídas em torno desse limite, respeitando esse limite e, talvez, até mesmo, aproveitando-se dele e o reforçando.

5.5.5.3 Hipótese III

A terceira e última hipótese que teremos o disparate de levantar diz respeito à provável influência da WM sobre a criação e formação das palavras de uma língua. Já mencionamos que consideramos que as línguas e a cognição humanas foram construídas em torno de uma limitação muito básica e muito estrita: a limitação de unidades independentes de informação manipuláveis pela WM. Acreditamos também que o cérebro humano se diferenciou — quantitativamente, como defende Herculano-Houzel (2012) — do cérebro dos demais animais não humanos em virtude e a reboque do surgimento e da evolução da linguagem, mas tal ideia não será desenvolvida neste trabalho.

Restringir-nos-emos, então, a tentar entender a influência da WM sobre certos aspectos fonológicos das línguas. Goldin-Meadow (2005) propõe que as crianças começam a aprender suas línguas pela prosódia (intrauterinamente), e seguem para as vogais (6 meses), consoantes (9 meses), palavras (10 a 15 meses), e daí para a morfologia e a sintaxe. O interessante nessa proposta é que a aquisição começaria “de fora para dentro”, do suprasegmentar para o segmentar, da prosódia para as sílabas fonológicas.

De certa forma, há um paralelo entre essa concepção de aquisição das línguas e a abordagem dos defensores da morfologia prosódica (*prosodic morphology*), que trabalham com a hipótese de que existe uma espécie de molde prosódico que limita o universo de possibilidades que poderia ser gerado a partir da simples justaposição de segmentos morfológicos independentes. Haveria uma hierarquia de segmentos prosódicos percorrendo a sequência mora-> sílaba-> pé -> palavra prosódica, e uma relação de binaridade entre segmentos inferiores e superiores: duas moras formariam uma sílaba; duas sílabas formariam um pé; e dois pés formariam uma palavra.

Mencionamos essa abordagem prosódica da formação das palavras porque a maioria dos estudos que identificamos na busca por evidências de que nossa hipótese III poderia fazer sentido analisavam os limites de tamanho de palavras e morfemas a partir da morfologia prosódica. Nossa hipótese III é que os morfemas – que definiremos como as menores unidades segmentares com conteúdo semântico – de qualquer língua não ultrapassam o limite de 3 sílabas fonológicas, e preferencialmente não ultrapassam o limite de 2 sílabas fonológicas. A justificativa para esse limite não seria a morfologia prosódica, mas a WM. Os limites de tamanho de palavras que a morfologia prosódica identificou e analisa não seriam consequência de fenômenos prosódicos universais, mas de limites da WM. E é bem possível que as restrições propostas pela morfologia prosódica estejam a serviço das limitações da WM, facilitando o gerenciamento desse limite por parte de quem está adquirindo uma língua (como no caso dos bebês, citado acima) ou, simplesmente, fazendo uso da língua. Sob esse ponto de vista, um “gesto prosódico” seria um auxiliar da WM, como ocorre, por exemplo, na segmentação prosódica dos números de telefone, de CPF etc.

A lógica por trás da hipótese III é que os aprendizes da língua manipulam sílabas fonológicas como unidades independentes, seguindo regras fonotáticas. Cada sílaba fonológica corresponderia a uma unidade independente no foco de atenção, e a WM conseguiria reter apenas cerca de 3 dessas sílabas ao mesmo tempo (talvez, apenas 2). Isolando e simplificando bastante o problema, e ignorando outros componentes do fenômeno, a ideia é que quando um aprendiz inocente, sem contato prévio com outras línguas, está aprendendo uma palavra e esta palavra contém mais de uma sílaba fonológica, ele precisa aglutinar as sílabas fonológicas em uma única unidade de informação. Por exemplo, digamos que o aprendiz encontra a palavra “aprendiz”. Digamos que ele já teve contato prévio com os sons da língua e já tem algum conhecimento de quais seriam suas sílabas fonológicas e como seria, grosso modo, sua fonotaxe. Ele, então, está defronte de uma palavra que, para ele, não é uma palavra, um item lexical, mas uma lista com 3 itens (3 sílabas fonológicas): “a”, “pren”, “diz”. A partir do momento em que ele associa esses 3 itens fonológicos ao significado de “aprendiz”, ocorre uma aglutinação em que

a lista composta por “a”+”pren”+”diz” passa a ser uma única unidade de informação, um item lexical com conteúdo semântico: “aprendiz”.

Esse processo seria aplicado a todas as associações de sílabas fonológicas com que o aprendiz se confrontasse. A cada novo confronto, uma nova sequência de sílabas seria aglutinada e armazenada na LTM como uma nova unidade de informação, associada a conteúdo semântico. Qualquer material fonológico minimamente associado a conteúdo semântico é, por definição, um morfema. Então, o aprendiz estaria, progressivamente, registrando novos morfemas em sua LTM. Entretanto, a construção de novos morfemas estaria sempre limitada pela capacidade da WM, que restringiria o número de sílabas fonológicas independentes, destituídas de sentido, que poderiam ser manipuladas e aglutinadas em uma unidade de informação.

Ocorre que, quando o aprendiz já tem registrado em sua LTM algum morfema, esse morfema passa a ser uma unidade de informação, mesmo que ele contenha 2 ou 3 sílabas fonológicas. Então, a palavra “aprendiz”, uma vez registrada na LTM, deixa de ser uma lista de sílabas e passa a ser uma única unidade. Dessa forma, quando o aprendiz encontra a palavra “aprendizagem”, que contém 5 sílabas fonológicas — e considerando que ele ainda não tenha tido contato prévio com o morfema/sufixo de derivação *-agem* —, ele não estará frente a uma lista de 5 sílabas fonológicas, 5 unidades de informação independentes, mas, sim, frente a uma lista com 3 unidades independentes, manipuláveis pela WM: “aprendiz”+”a”+”gem”. E, se ele aprender o morfema *-agem*, esse morfema passará a ser uma unidade de informação, facilitando, por exemplo, o aprendizado da palavra “bobagem”: “bo”+”bo”+”agem” (que, esperamos, não se aplique a esta hipótese).

A consequência das limitações impostas pela WM ao processo de aquisição das línguas é que todo e qualquer morfema, considerado como uma unidade mínima de sentido, terá, no máximo, 3 sílabas fonológicas. Qualquer palavra com um número maior de sílabas fonológicas já terá, inscrito em si, algum morfema que estará aglutinando sílabas fonológicas e reduzindo o número de unidades de informação independentes a apenas 2 ou 3.

Exceções a essa regra devem existir, e acreditamos que elas serão consequência do registro escrito das línguas. A escrita permite manter registrados itens lexicais que poderão ser aprendidos com esforço e dedicação. Qualquer um de nós seria capaz de incluir no léxico uma palavra com, digamos, 10 sílabas: *arfimetoleprigaterisca*. Bastaria ir decorando a palavra, aos poucos, até que ela se tornasse uma unidade na LTM (o que, particularmente, não recomendamos). Essa palavra não significa nada,

mas podemos atribuir-lhe algum sentido arbitrário a passar a usá-la no dia a dia. Entretanto, se não registrarmos a palavra no léxico escrito, acredito que haveria uma tendência para que ela desaparecesse ao longo da transmissão intergeracional da língua oral. Uma criança na fase inicial de aquisição da língua precisaria transformá-la em algo com apenas 2 ou 3 sílabas fonológicas. “— Filho, cuidado com aquela *arfimetoleprigaterisca!*” “—*Telisca? Onde?*” De alguma forma, essa criança teria registrado *telisca* como um morfema da língua. Ela poderia acrescentar a esse morfema, com o tempo, suas demais sílabas, mas o morfema *telisca* já estaria registrado na LTM. Dessa forma, não acreditamos que uma palavra como *arfimetoleprigaterisca* pudesse sobreviver muito tempo aos mecanismos de aquisição da língua por seus novos falantes, sem o apoio de um registro escrito que permitisse novas formas de aprendizagem.

Estudos de morfologia prosódica sugerem, por exemplo, um limite máximo de 2 sílabas para as raízes de línguas como o Maori (LACY, 2004), o Tcheco (KETNER, 2006) e o alemão (GOLSTON; WIESE, 1998); e uma análise nossa, rápida e superficial, do italiano e do português, aponta na mesma direção. A produção fonológica inicial de crianças parece, mesmo, corresponder a unidades CV (consoante vogal) e CVCV (FEE; AVERY, 1995). E Aitchison e Chiat (1981), em estudo mais antigo, observaram que erros fonológicos de crianças mais velhas (4, 6 e 9 anos) na aprendizagem de palavras novas assemelham-se aos erros fonológicos de aprendizes nas fases iniciais da aquisição da língua; e os autores alinham-se à nossa ideia (ou alinhamo-nos à ideia deles) de que a principal limitação para a ocorrência desses erros seria a memória: “*we have demonstrated that patterns of recall errors made by four- to nine-year olds in a word-learning task closely resemble certain phonological processes found in the speech of very young children. We suggest that the role of memory, which has hitherto been virtually ignored by linguists, needs to be taken into consideration*”¹ (AITCHISON; CHIAT, 1981, p. 323).

5.6 Sugestões para a evitar incongruências de fala lida

Este último item da dissertação busca atender ao último objetivo específico elencado na Introdução do trabalho: “elaborar, com base nos achados do estudo, possíveis recomendações para o

¹ Nós demonstramos que os padrões de erros de evocação cometidos por crianças com idades entre quatro e nove anos em uma tarefa de aprendizado de palavras novas é muito semelhante a certos processos fonológicos encontrados na fala de crianças bem mais novas. Nós sugerimos que o papel da memória, que vem sendo virtualmente ignorado pelos linguistas, precisa ser levado em consideração. (tradução nossa)

aperfeiçoamento do processo de produção de textos escritos cujo objetivo seja a leitura e voz alta, especialmente quando relacionada ao fonostilo político”.

O impulso inicial para a realização deste trabalho foi encontrar meios de reduzir as incongruências durante a leitura em voz alta. No caminho, acabamos nos descolando desse objetivo e passeando por outras veredas nos caminhos da linguística. Entretanto, apesar de não se tratar mais do ponto de chegada do trabalho, acreditamos que determinadas incongruências poderiam ter sido evitadas caso certas adaptações tivessem sido feitas nos textos escritos. Não se trata, obviamente, de uma prescrição peremptória, baseada em evidências de grande qualidade, pois, para isso, seriam necessários experimentos controlados; mas, de qualquer forma, registramos algumas sugestões que futuros estudos poderão, eventualmente, validar ou negar.

Nossa ideia preliminar era identificar características do texto escrito que pudessem contribuir para a ocorrência de incongruências. Não nos propusemos a lidar com características atribuíveis aos leitores, pois, além de muito variadas, implicariam abordagens metodologicamente e eticamente diferentes das que utilizamos. No entanto, não podemos deixar de fazer algumas observações genéricas sobre a relação entre as incongruências e a forma como o leitor enfrenta o texto escrito.

Em primeiro lugar, como já discutimos, parece haver uma correlação entre incongruência e oralidade: quanto mais o leitor tenta parecer espontâneo; quanto mais ele tenta aproximar a fala lida da fala espontânea; e quanto mais ele tenta dar expressividade à fala lida, mais evidentes ficam as eventuais incongruências que ele produz. É possível que estejamos lidando com uma simples questão de contraste: certas incongruências não caberiam na fala espontânea e, portanto, ficariam mais evidentes quando a fala lida tentasse se aproximar dela.

Em segundo lugar, e em complemento à observação anterior, percebemos que uma fidelidade absoluta ao texto escrito não parece ser salutar. Oradores que usam o texto escrito apenas como um guia geral, recriando partes do texto com palavras e construções linguísticas próprias, produzem menos incongruências do que os que tentam ler cada vírgula do texto escrito. Trata-se, naturalmente, de uma correlação, não de uma relação de causa e efeito. Pode ser que os oradores que improvisam e utilizam o texto apenas como um guia geral sejam mais felizes na leitura porque leem (silenciosamente) mais segmentos dos períodos antes de produzir sua fala. Pode ser que eles sejam mais capazes de ler mais segmentos, ou, simplesmente, que eles tenham de fazer isso para conseguir improvisar. Neste último caso, a maior congruência seria um efeito colateral da tentativa de improvisar, e não uma consequência

direta de uma capacidade de ler mais material linguístico antes de falar. De qualquer forma, parece que não é bom ater-se fielmente ao que está escrito e isso seria mais um indicativo de que a escrita (como a conhecemos) não é um bom guia para a reprodução da fala.

No entanto, como dependemos da escrita para a fala lida, tentaremos identificar algumas sugestões de “aperfeiçoamentos” para reduzir a incidência das incongruências que observamos. Separamos as sugestões em dois grupos. O primeiro tratará das incongruências segmentares, principalmente das mais comuns: desvios de ortoépia e troca de palavras. O segundo, das incongruências suprasegmentares, especialmente as inserções/deleções de fronteiras prosódicas. A repartição das sugestões em dois grupos não significa que cada sugestão seja útil apenas para o seu grupo de incongruências, mas, simplesmente, que julgamos que ela teria maior potencial naquela situação. Em tese, qualquer característica do texto escrito que consiga reduzir qualquer tipo de incongruência poderia produzir efeitos benéficos para a congruência da leitura como um todo.

5.6.1 Incongruências segmentares

Como comentamos, para não produzir incongruências segmentares, basta ler o que está escrito. Qualquer desvio com relação ao que está escrito seria, em princípio, completamente evitável. As sugestões para evitar esses desvios são, via de regra, bastante óbvias:

- Usar fontes com tamanho e formatação adequados às preferências e necessidades individuais de cada leitor
- Evitar translineações que fragmentem palavras (e isso serve também para evitar incongruências suprasegmentares, como as pausas e lentificações)
- Pode ser que o uso de colunas, com parágrafos mais estreitos, com menos palavras por linha (como mencionamos anteriormente) seja benéfico
- Devem-se evitar palavras e associações de palavras que ofereçam alguma dificuldade articulatória, como as que escrevemos acima: “prescrição peremptória”. Também observamos dificuldades onde havia repetição de fonemas semelhantes: noBRE PROpósito, DE DESestimular, DE DESocupação, AO AUxílio, AO AUtoritarismo

5.6.2 Incongruências suprasegmentares

Como já discutimos, as incongruências suprasegmentares caracterizadas por inserção/deleção de fronteira prosódica parecem “revelar que o falante já tomou determinadas decisões com relação à

organização do período e que tais decisões não corresponderam à realidade linguística pretendida pelo texto escrito”. Portanto, parece necessário ajuda-lo a tomar as decisões corretas; mostrar-lhe os limites e a organização dos períodos; fornecer alguma indicação antecipada do que está para acontecer, para que ele não tenha de tomar decisões sem saber o que vem pela frente. Seria como sinalizar uma estrada para que o motorista consiga evitar acidentes: rua sem saída, proibido virar à esquerda, não parar, curva sinuosa adiante... Identificamos que as relações hierárquicas desniveladas estavam associadas a incongruências de fronteiras e que os falantes têm uma tendência para inserir fronteiras. Acreditamos que isso não aconteceria se os falantes pudessem “enxergar” os limites do período e de seus constituintes.

Garcia (2011), em seu manual sobre como “escrever bem” em PB, levanta uma questão interessante. Ele classifica os períodos em “frouxos” ou “tensos”, e contraindica o uso dos primeiros. Períodos frouxos seriam aqueles que, em nossa análise, apresentavam desnivelamento hierárquico após a predicação primária. Por exemplo, no provérbio modificado “todos os gatos são pardos de noite”, construiu-se um período frouxo (na classificação de Othon M. Garcia), em que a predicação primária — todos os gatos são pardos — poderia induzir a inserção de uma fronteira terminal, já que o constituinte adverbial não é obrigatório. Garcia (2011), obviamente, não levou em conta a nossa abordagem, mas disse que, via de regra, os provérbios empregam uma estratégia de “prótase” seguida por “apódose”, ou condicionante seguido por condicionado. Por exemplo: a) de noite, todos os gatos são pardos; b) de hora em hora, Deus melhora; e assim por diante.

Nas palavras de Garcia (2011, p. 72), em períodos frouxos, “como o essencial já foi dito, o secundário torna-se (...) quase desprezível, sendo bem provável que o leitor ‘passe por cima’”. De certa forma, é o que acontece quando o leitor insere uma fronteira onde não havia, principalmente quando essa fronteira é terminal. Em períodos tensos, por outro lado, não seria possível encerrar o período ou o constituinte antes da hora. Nesse trecho de Raul Pompeia (Apud GARCIA; 2011, p. 74), que apresenta o que Garcia (2011) chama de “prótese atenuada”, essa característica também fica clara: “o homem, por desejo de nutrição e de amor, produziu a evolução histórica da humanidade”. Especialmente quando confrontado com sua versão “frouxa”: o homem produziu a evolução histórica da humanidade por desejo de nutrição e de amor.

A abordagem de Garcia (2011) é de caráter mais estilístico, apesar de podermos argumentar que estilo e cognição devem estar, de alguma forma, associados. O uso de períodos “tensos” — ou, pelo menos, o cuidado de construí-los o menos frouxos possível — faria com que houvesse menos

desnivelamentos hierárquicos após a predicação primária e isso poderia inibir a inserção inadequada de fronteiras. Esta é, portanto, uma das sugestões que podemos fazer. Entretanto, também podemos pensar em sugestões que não interfiram, diretamente, no conteúdo e na forma das construções linguísticas, já que não são os períodos “frouxos” ou com desnivelamentos que produzem incongruências, mas, sim, os falantes. Voltamos, então, à proposta de sinalizar melhor o que está sendo lido.

O papel de sinalizador do material escrito é atribuído, tradicionalmente, ao sistema de pontuação. Krahn (2014), em tese de doutorado, investiga a origem da pontuação na escrita e seu uso corrente na língua inglesa, e traz algumas informações interessantes. Ele diz que, historicamente, a pontuação era usada como um auxílio para a leitura em voz alta e que seu uso como sinalizador para a leitura silenciosa surgiu apenas posteriormente. Acrescenta que hoje, na língua inglesa, ela é considerada ou como dispositivo retórico, ou como marcador gramático, mas que ele, após seus estudos, concluiu que a pontuação serve, primordialmente, para demarcar constituintes e evitar confusões por parte do leitor. Textualmente, ele diz que: “*the goal of dynamic (horizontal) punctuation is to preserve the integrity of the canonical (SVO) sentence and make it visible by using commas to prevent inappropriate modification*”¹ (KRAHN, 2019, p. 167).

Vemo-nos aqui, então, num retorno às origens da pontuação: auxiliar a leitura em voz alta. É certo que a pontuação, por si só, da forma como tem sido utilizada, não é suficiente (em PB, pelo menos) para evitar incongruências. Talvez por ser, segundo Raso (2013, p. 23), “um sistema de marcação insuficiente e enganador”. Ou, talvez, por ser muito discreta, pequena, pouco intrusiva. Senão, vejamos: “.” “;” (ponto final e vírgula) são, de fato, bastante sutis. Além de poderem, eventualmente, passar despercebidos durante a leitura, tornam-se ainda menos perceptíveis à distância, quando estão adiante do ponto focal da visão.

Uma possível solução para isso seria tornar a pontuação menos discreta:

¹ O objetivo da pontuação dinâmica (horizontal) é preservar a integridade dos períodos canônicos (SVO) e torná-la visível pelo uso de vírgulas, evitando alterações inapropriadas. (tradução nossa)

Poderíamos , então , empregar uma pontuação mais “vigorosa” , como esta ●

Existe a possibilidade de que esse tipo de artifício reduza as inserções/deleções incongruentes de fronteira prosódica, mas acreditamos que há soluções menos duvidosas e com maior potencial de sucesso. Harrison (2007), por exemplo, trabalhando com estudantes de latim, propôs estruturar o texto separando seus constituintes, frases e/ou orações em linhas diferentes, com o objetivo de facilitar sua compreensão pelos alunos. Na frase “*Dissertatio de methodo recte utendi ratione et veritatem in scientiis investigandi*”¹, para evitar que o leitor delete a fronteira entre *ratione* e *veritatem*, o autor propôs separar os constituintes em linhas diferentes:

Dissertatio de methodo
recte utendi ratione
et veritatem in scientiis investigandi

O objetivo seria evidenciar as duas estruturas chamadas “gerundivos”, qualificadoras de “*methodo*” (*utendi* e *investigandi*), e evitar que a presença do “*et*” induza o leitor a ler “*recte utendi ratione et veritatem*”. A necessidade por trás dessa proposta é bastante parecida com a nossa necessidade: fazer com que o leitor enxergue, num passar de olhos, a estrutura geral do período — onde começa, como se organiza, e onde termina.

Além disso, se o leitor tiver uma ideia melhor da organização do período, torna-se mais fácil para ele afastar-se do texto escrito e improvisar, obtendo os benefícios que já mencionamos. Na verdade, as limitações impostas por um sistema muito rígido de representação gráfica da fala, com uma padronização muito estrita, devem ter dificultado a evolução dessa representação. Quando os usuários da língua têm mais liberdade, seus processos evolutivos naturais podem continuar agindo e, via de regra, tornando a língua (incluindo seu registro gráfico) mais ergonômica para a cognição. É o que observamos, por exemplo, na forma como os jovens (e não tão jovens) costumam escrever no

¹ Discurso sobre o método de como bem conduzir a razão e buscar a verdade na ciência. (tradução nossa)

aplicativo de troca de mensagens Whatsapp, onde é comum recebermos mensagens bastante fragmentadas. Num caso concreto, em vez de se escrever:

“Oi, André. Olhei aqui, não atrasou, era pra hoje. Aí, tem um dia pra revisão. Então, poderia voltar pra você na segunda-feira. Ou seja, o prazo seria pra segunda, para todo o processo. A data que aparece pra você é a sua data de entrega.”

Escreveu-se:

“Oi André
Olhei aqui
Não atrasou
Era pra hoje
Aí tem um dia pra revisão
Então poderia voltar pra você na segunda-feira
Ou seja, o prazo seria pra segunda
Pra todo o processo
A data que aparece pra você é a sua data de entrega”

Fragmentou-se, então, toda a mensagem. Cada oração, ou conjunto de orações, ou mesmo adjuntos, virou uma mensagem isolada, uma nova linha. Ao invés de construir um parágrafo que contivesse a ideia como um todo, parece que cada fragmento da ideia foi sendo transformado em texto à medida que ia surgindo na mente, como costuma ocorrer na fala espontânea. Outro exemplo, em que se responde à pergunta “onde você está?”, ficou assim:

Ué
Aqui
Onde me deixou

Essa “escrita de WhatsApp” está em franco contraste com a escrita padrão, em que nos esforçamos para reunir e organizar e aglutinar nossa mensagem em períodos e parágrafos encadeados, fluidos, menos “entrecortados”. É possível que nosso esforço em busca de fluidez seja contraproducente quando o objetivo é escrever textos para serem lidos em voz alta. Ou que, talvez, seja necessário indicar melhor para o leitor os contornos e limites dessa fluidez. O objetivo, afinal, é produzir uma escrita que seja mais suficiente e menos enganadora.

O que sugerimos, então — mais como hipótese de trabalho do que como recomendação —, é uma escrita que atenda, minimamente, aos seguintes parâmetros:

- Cada período (enunciado) deve estar visualmente agrupado, separado dos demais períodos, e não pode ser graficamente fragmentado
- Fronteiras prosódicas não terminais serão representadas graficamente por uma quebra de linha (mas nem toda quebra de linha precisa representar uma fronteira prosódica)

Consideremos o segmento de texto abaixo, retirado do corpus:

“Prevê-se que o aumento da renda real do trabalhador brasileiro neste ano chegue a 2,5%, depois de dois anos consecutivos de perdas. Outro dado econômico bastante alentador diz respeito ao emprego.”

Seguindo os dois parâmetros acima, os períodos desse texto seriam representados graficamente assim:

Prevê-se
que o aumento da renda real do trabalhador brasileiro
neste ano
chegue a 2,5%,
depois de dois anos consecutivos de perdas.

Outro dado econômico
bastante alentador
diz respeito
ao emprego.

Poder-se-ia fragmentar menos o texto? Sim. Poder-se-ia fragmentá-lo mais? Sim. Poder-se-ia fragmentá-lo de forma diferente? Sim. A ideia geral (além de evitar as mesóclises) é evitar a falsa sensação de segurança que o texto contínuo transmite; evidenciar os limites de cada período; e reduzir as chances de inserção ou deleção incongruente de fronteiras, tanto terminais, quanto não terminais. Podem-se utilizar mais colunas na mesma página, evitando o uso excessivo de novas páginas para acomodar um texto verticalmente mais alongado. Podem-se também usar novas tecnologias, como telas eletrônicas, evitando o gasto de papel e a quebra de páginas. E pode acontecer, também, que essas

sugestões não tragam benefício algum para os leitores. Como já repetimos algumas vezes nesse trabalho, apenas um estudo experimental sanaria essas dúvidas.

5.6.3 Sugestão adicional (e final): sobre a escrita de números

Havia em nosso corpus 226 números escritos com algarismos. Eram datas (ex.:7 de setembro de 1922), anos (ex.:2017), porcentagens (ex.: 23%) e numerais cardinais de um modo geral. Em 8% dos casos, houve incongruências. Foram 6 casos de desvio de ortoépia, 4 casos de troca de um numeral por outro, 6 pausas durante a leitura do numeral, e 2 hesitações. Esses resultados não foram descritos no capítulo homônimo porque os julgamos secundários. De qualquer forma, talvez seja útil que numerais mais complexos, maiores, na casa das centenas, milhares etc., sejam escritos por extenso. Por exemplo: R\$ 3.579.462,31 poderia ser escrito como se faziam nos ultrapassados cheques: três milhões, quinhentos e setenta e nove mil, quatrocentos e sessenta e dois reais e trinta e um centavos. Como já comentamos, essa recomendação só poderia ser levada ao pé da letra se fizéssemos um estudo controlado, comparando a congruência de leitura de numerais escritos por extenso *versus* a congruência de leitura de numerais escritos como algarismos.

6. CONCLUSÃO

Uma conclusão inevitável deste trabalho é que devemos, sempre, ajustar nossos objetivos às nossas possibilidades. E, de preferência, ajustá-los bem abaixo de nossas possibilidades, porque, via de regra, costumamos superestimar nossas capacidades e o tempo de que dispomos. Trata-se de um viés cognitivo que tem até um nome: *planning fallacy*, falácia do planejamento. Em termos práticos, parece ser preferível delimitar muito bem o objeto de estudo e, se possível, ampliá-lo, ao invés de não o delimitar e ter, em algum momento, de abandonar objetivos.

Em nosso caso, estabelecemos objetivos que se revelaram muito mais distantes do que prevíamos. Durante a realização do estudo, fomos descobrindo que as possibilidades teóricas relacionadas aos problemas que pretendíamos estudar eram muito mais largas, profundas e revoltas do que imaginávamos. Não abandonamos nossos objetivos iniciais, mas confessamos que foi muito mais difícil atingi-los ou nos aproximarmos deles do que julgamos em nosso falacioso planejamento inicial. Em nosso favor, podemos dizer que fomos persistentes. E, em nosso desfavor, que fomos bastante ingênuos, para dizer o mínimo.

As conclusões seguintes já foram, de certa forma, mencionadas ao longo dos resultados e das discussões, mas as consolidamos aqui. Em primeiro lugar, a leitura congruente em voz alta, com aderência completa ao texto escrito, parece ser uma obra de ficção, uma miragem que desaparece rapidamente e que revela, em seu lugar, uma realidade linguística alternativa. Parece que há, entre a escrita e a fala, muito mais obstáculos e diferenças do que supõe nossa vã filosofia linguística. O fato é que as incongruências de leitura em voz alta — como desvios de ortoépia, trocas de palavras, pausas e inserções de fronteiras prosódicas — foram extremamente comuns em nosso corpus e não pouparam nenhum de nossos 39 participantes. Além disso, observamos que as retratações (ou não retratações) das incongruências também ofereceram um campo muito fértil para estudos linguísticos.

Parece haver associações complexas entre os textos escritos e as incongruências associadas a eles. No caso das trocas de palavras, além sofrerem influência das categorias gramaticais, da frequência de cada palavra na língua e da semelhança fonológica, há palavras com enorme resistência às trocas, como a palavra “que”. No caso das incongruências de fronteiras prosódicas, notamos que há grande tendência à inserção incongruente de fronteiras e que, surpreendentemente, não costuma haver retratação desse tipo de incongruência.

Considerando tudo o que observamos ao longo do trabalho, levantamos (ou reeditamos) algumas hipóteses que merecem sofrer o escrutínio e a crítica de verdadeiros linguistas. Entre elas, mencionamos as seguintes. A fala (externa ou interna) só é compreendida e monitorada após ser “articulada”, produzida. As fronteiras prosódicas, via de regra, não são produzidas para facilitar a compreensão do ouvinte, mas, sim, para permitir a compreensão do próprio falante (ao contrário da função prosódica de marcar proeminência, que parece ser direcionada ao ouvinte). A memória de trabalho merece receber mais atenção de linguistas e cientistas cognitivos, pois parece ser um dos elementos mais importantes a limitar e moldar as possibilidades das engenharias linguística e cognitiva. No caso específico da linguagem, acreditamos que as limitações da memória de trabalho são responsáveis pela tendência que observamos de inserir fronteiras prosódicas incongruentes durante a leitura em voz alta. Além disso, levantamos as hipóteses de que a memória de trabalho limita o número de participantes obrigatórios de um evento verbal (estrutura argumental) a apenas dois participantes; e que também limita o número de sílabas fonológicas dos morfemas de qualquer língua a, no máximo, três.

Independentemente de nossas hipóteses conterem alguma semente da verdade ou serem completamente despropositadas, relembremos Lévi-Strauss: "teorias vivem pouco, dados e descrições nunca morrem". Nossas hipóteses são apenas embriões (talvez anencéfalos) de teorias, mas nossos dados e descrições são o que são: dados e descrições, obtidos com a melhor intenção de fidelidade à realidade fática em que foram produzidos. Se deste trabalho nada mais se aproveitar, que se aproveitem, ao menos, os dados e as descrições.

7. REFERÊNCIAS

ABAURRE, Maria Bernadete Marques; SILVA, Ademar da. O desenvolvimento de critérios de segmentação na escrita. **Temas em Psicologia**, Ribeirão Preto, Sp, v. 1, n. 1, p.89-102, jan. 1993.

AITCHISON, Jean; CHIAT, Shulamuth. Natural Phonology or Natural Memory? the Interaction Between Phonological Processes and Recall Mechanisms. **Language And Speech**, [S.L.], v. 24, n. 4, p. 311-326, out. 1981. SAGE Publications. <http://dx.doi.org/10.1177/002383098102400402>.

ALBOUKADEL KASSAMBARA. Datanovia (ed.). **Cohen's Kappa in R: For Two Categorical Variables**. 2021. Disponível em: <https://www.datanovia.com/en/lessons/cohens-kappa-in-r-for-two-categorical-variables/#:~:text=Cohen%27s%20Kappa%20in%20R%3A%20For%20Two%20Categorical%20Variables,-20%20mins&text=This%20process%20of%20measuring%20the,is%20called%20inter%2Drater%20reliability.&text=The%20Cohen%27s%20kappa%20is%20a,that%20removes%20this%20chance%20agreement>. Acesso em: 24 fev. 2021.

ALDAY, Phillip M.; SCHLESEWSKY, Matthias; BORNKESSEL-SCHLESEWSKY, Ina. Towards a Computational Model of Actor-Based Language Comprehension. **Neuroinformatics**, [S.L.], v. 12, n. 1, p. 143-179, 4 ago. 2013. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s12021-013-9198-x>.

ALDERSON-DAY, Ben; FERNYHOUGH, Charles. Inner speech: development, cognitive functions, phenomenology, and neurobiology.. **Psychological Bulletin**, [S.L.], v. 141, n. 5, p. 931-965, set. 2015. American Psychological Association (APA). <http://dx.doi.org/10.1037/bul0000021>.

ALTMANN, Gerry T.m.; KAMIDE, Yuki. Incremental interpretation at verbs: restricting the domain of subsequent reference. **Cognition**, [s.i.], v. 73, n. 0, p.247-264, set. 1999.

AYTO, John (Ed.). garden: lead someone up the garden path. In: AYTO, John (Ed.). **Oxford Dictionary of English Idioms**. 3. ed. [s.i.]: Oxford University Press, 2009. p. 143.

BADDELEY, Alan. Working memory. **Science**, [s.l.], v. 255, n. 5044, p.556-559, 31 jan. 1992. American Association for the Advancement of Science (AAAS). <http://dx.doi.org/10.1126/science.1736359>.

BARBOSA, Plínio A.. Conhecendo melhor a prosódia: aspectos teóricos e metodológicos daquilo que molda nossa enunciação. **Revista de Estudos da Linguagem**, Belo Horizonte, Mg, v. 20, n. 1, p.11-27, jun. 2012.

BARBOSA, Plínio A.; RASO, Tommaso. A segmentação da fala espontânea: aspectos prosódicos, funcionais e aplicações para a tecnologia / Spontaneous Speech Segmentation. **Revista de Estudos da Linguagem**, [s.l.], v. 26, n. 4, p.1397-1433, 11 out. 2018. Faculdade de Letras da UFMG. <http://dx.doi.org/10.17851/2237-2083.26.4.1397-1433>.

BARBOSA, Plinio A.. Prosódia. São Paulo (sp): Parábola Editorial, 2019. 133 p. (Linguística para o ensino superior).

BASILIO, Margarida. **Formação e classes de palavras no português do brasil**. 3. ed. São Paulo (Sp): Contexto, 2004. 87 p.

BEYMER, David; RUSSELL, Daniel M.; ORTON, Peter Z.. Wide vs. Narrow Paragraphs: an eye tracking analysis. **Human-Computer Interaction - Interact 2005**, [S.L.], p. 741-752, 2005. Springer Berlin Heidelberg. http://dx.doi.org/10.1007/11555261_59.

BICKEL, Balthasar. Grammatical Relations Typology. **Oxford Handbooks Online**, [s.l.], p.1-41, 25 nov. 2010. Oxford University Press. <http://dx.doi.org/10.1093/oxfordhb/9780199281251.013.0020>.

BLUMENTHAL-DRAMÉ, Alice; MALAIA, Evie. Shared neural and cognitive mechanisms in action and language: The multiscale information transfer framework. **Wiley Interdisciplinary Reviews: Cognitive Science**, [s.l.], v. 10, n. 2, p.1-18, 12 nov. 2018. Wiley. <http://dx.doi.org/10.1002/wcs.1484>.

BÖGELS, Sara et al. Prosodic Breaks in Sentence Processing Investigated by Event-Related Potentials. **Language And Linguistics Compass**, [s.l.], v. 5, n. 7, p.424-440, jul. 2011. Wiley. <http://dx.doi.org/10.1111/j.1749-818x.2011.00291.x>.

BORNKESSEL-SCHLESEWSKY, Ina; SCHLESEWSKY, Matthias. The Role of Prominence Information in the Real-Time Comprehension of Transitive Constructions: A Cross-Linguistic Approach. **Language And Linguistics Compass**, [s.l.], v. 3, n. 1, p.19-58, jan. 2009. Wiley. <http://dx.doi.org/10.1111/j.1749-818x.2008.00099.x>.

BORNKESSEL-SCHLESEWSKY, Ina et al. Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. **Brain And Language**, [s.l.], v. 117, n. 3, p.133-152, jun. 2011. Elsevier BV. <http://dx.doi.org/10.1016/j.bandl.2010.09.010>.

BORNKESSEL-SCHLESEWSKY, Ina; SCHLESEWSKI, Mathias. Competition in argument interpretation: evidence from the neurobiology of language. In: MACWHINNEY, Brian; MALCHUKOV, Andrej; MORAVCSIK, Edith (Ed.). **Competing Motivations in Grammar and Usage**. New York, Ny: Oxford University Press, 2014. p. 107-126.

BORNKESSEL-SCHLESEWSKY, Ina; STAUB, Adrian; SCHLESEWSKY, Matthias. The Timecourse of Sentence Processing in the Brain. **Neurobiology Of Language**, [s.l.], p.607-620, 2016. Elsevier. <http://dx.doi.org/10.1016/b978-0-12-407794-2.00049-3>.

BORTFELD, Heather; LEON, Silvia D.; BLOOM, Jonathan E.; SCHOBER, Michael F.; BRENNAN, Susan E.. Disfluency Rates in Conversation: effects of age, relationship, topic, role, and gender. **Language And Speech**, [S.L.], v. 44, n. 2, p. 123-147, jun. 2001. SAGE Publications. <http://dx.doi.org/10.1177/00238309010440020101>.

BRANDÃO, Silvia Figueiredo; VIEIRA, Silvia Rodrigues. Concordância nominal e verbal: contribuições para o debate sobre o estatuto da variação em três variedades urbanas do português. **Alfa: Revista de Linguística** (São José do Rio Preto), [S.L.], v. 56, n. 3, p. 1035-1064, 2012. FapUNIFESP (SciELO). <http://dx.doi.org/10.1590/s1981-57942012000300013>.

BRASIL. Senado Federal. Senado Federal. **Senado Multimídia**: Eventos Legislativos. 2019. Disponível em: <<https://www12.senado.leg.br/multimidia>> . Acesso em: 17 nov. 2019.

BRASIL. SENADO FEDERAL. **Atividade legislativa**: pronunciamentos. 2019. Disponível em: <<https://www25.senado.leg.br/web/atividade/pronunciamentos>> . Acesso em: 03 dez. 2019.

CALAIS, Lucila Leal; LIMA-GREGIO, Aveliny Mantovan; ARANTES, Pablo; GIL, Daniela; BORGES, Alda Christina Lopes de Carvalho. Um estudo sobre associação semântica de palavras do

português brasileiro. **Codas**, [S.L.], v. 28, n. 5, p. 497-506, 31 out. 2016. FapUNIFESP (SciELO). <http://dx.doi.org/10.1590/2317-1782/20162015268>.

CAMACHO, Roberto Gomes. Construções relativas nas variedades do português: uma interpretação discursivo-funcional. **Filologia e Linguística Portuguesa**, [S.L.], v. 15, n. 1, p. 179, 30 jun. 2013. Universidade de Sao Paulo, Agencia USP de Gestao da Informacao Academica (AGUIA). <http://dx.doi.org/10.11606/issn.2176-9419.v15i1p179-214>.

CAMARA JÚNIOR, Joaquim Mattoso. **Estrutura da língua portuguesa**. 32. ed. Petrópolis, Rj: Vozes, 2000. P. 69.

CAPLAN, David. Working Memory and Sentence Comprehension. **Neurobiology Of Language**, [s.l.], p.633-645, 2016. Elsevier. <http://dx.doi.org/10.1016/b978-0-12-407794-2.00051-1>.

CARNIE, Andrew. **Syntax: a generative introduction**. 3. ed. [s.i.]: Wiley-blackwell, 2013.

CARNIE, Andrew. **About Andrew Carnie**. 2019. Disponível em: <<https://linguistics.arizona.edu/user/andrew-carnie>> . Acesso em: 27 nov. 2019.

CAVALCANTE, Frederico; SILVA, Cassiano César da; RASO, Tommaso. Critérios de segmentação da fala. **Anais da XII Sevfale**, Belo Horizonte (Mg), p. 129-142, 1 jan. 2015. Disponível em: <http://www.anais.letras.ufmg.br/index.php/SEVFALE/XIISEVFALE/paper/view/18/14>. Acesso em: 24 fev. 2021.

CHACÓN, Dustin Alfonso; MOMMA, Shota; PHILLIPS, Colin. Linguistic representations and memory architectures: the devil is in the details. **Behavioral And Brain Sciences**, [S.L.], v. 0, n. 0, p. 24-25, 2016. Cambridge University Press (CUP). <http://dx.doi.org/10.1017/s0140525x15000746>.

CHOMSKY, Noam. Minimal Recursion: exploring the prospects. **Recursion: Complexity in Cognition**, [S.L.], p. 1-15, 2014. Springer International Publishing. http://dx.doi.org/10.1007/978-3-319-05086-7_1.

CHRISTIANSEN, Morten H.; CHATER, Nick. The Now-or-Never bottleneck: A fundamental constraint on language. **Behavioral And Brain Sciences**, [s.l.], v. 39, p.1-72, 14 abr. 2015. Cambridge University Press (CUP). <http://dx.doi.org/10.1017/s0140525x1500031x>.

COP, Uschi; DIRIX, Nicolas; DRIEGHE, Denis; DUYCK, Wouter. Presenting GECO: an eyetracking corpus of monolingual and bilingual sentence reading. **Behavior Research Methods**, [S.L.], v. 49, n. 2, p. 602-615, 18 maio 2016. Springer Science and Business Media LLC. <http://dx.doi.org/10.3758/s13428-016-0734-0>.

CORBALLIS, Michael C.. **The Recursive Mind**: the origins of human language, thought, and civilization. Princeton: Princeton University Press, 2014. 291 p.

COUPÉ, Christophe; OH, Yoon; DEDIU, Dan; PELLEGRINO, François. Different languages, similar encoding efficiency: comparable information rates across the human communicative niche. **Science Advances**, [S.L.], v. 5, n. 9, p. 1-10, set. 2019. American Association for the Advancement of Science (AAAS). <http://dx.doi.org/10.1126/sciadv.aaw2594>.

COLTHEART, Max. Modeling Reading: the dual-route approach. **The Science Of Reading: A Handbook**, [S.L.], p. 6-23, 2005. Blackwell Publishing Ltd. <http://dx.doi.org/10.1002/9780470757642.ch1>.

COWAN, Nelson. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. **Behavioral And Brain Sciences**, [S.L.], v. 24, n. 1, p. 87-114, fev. 2001. Cambridge University Press (CUP). <http://dx.doi.org/10.1017/s0140525x01003922>.

COWAN, Nelson; MOREY, Candice C.; NAVEH-BENJAMIN, Moshe. An Embedded-Processes Approach to Working Memory. In: LOGIE, Robert; CAMOS, Valérie; COWAN, Nelson. **Working Memory**: state of the science. Oxford: Oxford University Press, 2021. Cap. 3. p. 44-84.

CREISSELS, Denis. Cross-Linguistic Variation in the Treatment of Beneficiaries and the Argument vs. Adjunct Distinction. **Linguistic Discovery**, [S.L.], v. 12, n. 2, p. 41-55, 2014. Dartmouth College Library Press. <http://dx.doi.org/10.1349/ps1.1537-0852.a.445>.

CUNHA, Celso Ferreira da; CINTRA, Luis Filipe Lindley. **Nova gramática do português contemporâneo**. 3. ed. Rio de Janeiro: Nova Fronteira, 2001.

CYRINO, Sonia; ESPINAL, M. Teresa. Bare Nominals in Brazilian Portuguese: more on the dp/np analysis. **Natural Language & Linguistic Theory**, [S.L.], v. 33, n. 2, p. 471-521, 24 set. 2014. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s11049-014-9264-6>.

DAANE, Mary C. et al. **Fourth-Grade Students Reading Aloud**: NAEP 2002 Special Study of Oral Reading. Washington, Dc: National Center For Education Statistics, Institute Of Education Sciences, U.s. Department Of Education, 2005. 72 p. (The Nation's Report Card). Disponível em: <<https://nces.ed.gov/nationsreportcard/pdf/studies/2006469.pdf>> . Acesso em: 24 out. 2018.

DARCY RIBEIRO. [s.i.]: Tv Cultura, 1995. Son., color. Série Roda Viva. Disponível em: <<https://www.youtube.com/watch?v=AAFzOemlAbg>> . Acesso em: 02 dez. 2019.

DEHAENE, Stanislas. **Reading in the brain**: the science and evolution of a human invention. [s.i.]: Penguin Books, 2009.

DEHAENE, Stanislas. **Consciousness and the brain**: deciphering how the brain codes our thoughts. New York, Ny: Penguin Books, 2014.

DOWTY, David. Thematic proto-roles and argument selection. **Language**, [s.l.], v. 67, n. 3, p.547-619, 1991. Project Muse. <http://dx.doi.org/10.1353/lan.1991.0021>.

DUARTE, Paulo Mozânio Teixeira. **Do lexema e termos afins na terminologia gramatical**. 2004. Disponível em: http://www.repositorio.ufc.br/bitstream/riufc/26297/1/2004_art_pmtduarte.pdf. Acesso em: 30 abr. 2021.

ERICSSON, K. Anders; KINTSCH, Walter. Long-term working memory. **Psychological Review**, [s.l.], v. 102, n. 2, p.211-245, 1995. American Psychological Association (APA). <http://dx.doi.org/10.1037/0033-295x.102.2.211>.

EVANS, Nicholas; LEVINSON, Stephen C.. The myth of language universals: Language diversity and its importance for cognitive science. **Behavioral And Brain Sciences**, [s.l.], v. 32, n. 5, p.429-448, out. 2009. Cambridge University Press (CUP). <http://dx.doi.org/10.1017/s0140525x0999094x>.

FEE, E. Jane; AVERY, Peter. Segments and Syllables in Early Language Acquisition. In: ARCHIBALD, John. **Phonological Acquisition and Phonological Theory**. Hillsdale (Nj): Lawrence Earlbaum Associates, 1995. p. 45-46.

FERREIRA, Fernanda; CLIFTON JUNIOR, Charles. The Independence of Syntactic Processing. **Journal Of Memory And Language**, Massachusetts, Ma, v. 25, n. 0, p.348-368, jan. 1986.

FERREIRA, Fernanda; PATSON, Nikole D.. The Good Enough Approach to Language Comprehension. **Language And Linguistics Compass**, [s.l.], v. 1, n. 1-2, p.71-83, mar. 2007. Wiley. <http://dx.doi.org/10.1111/j.1749-818x.2007.00007.x>.

FERREIRA, Fernanda; ÇOKAL, Derya. Sentence Processing. **Neurobiology Of Language**, [s.l.], p.265-274, 2016. Elsevier. <http://dx.doi.org/10.1016/b978-0-12-407794-2.00022-5>.

FERREIRA, Fernanda; CHANTAVARIN, Suphasiree. Integration and Prediction in Language Processing: A Synthesis of Old and New. **Current Directions In Psychological Science**, [s.l.], v. 27, n. 6, p.443-448, 31 out. 2018. SAGE Publications. <http://dx.doi.org/10.1177/0963721418794491>.

FRAZIER, Lyn; CLIFTON, Charles; CARLSON, Katy. Don't break, or do. **Lingua**, [s.l.], v. 114, n. 1, p.3-27, jan. 2004. Elsevier BV. [http://dx.doi.org/10.1016/s0024-3841\(03\)00044-5](http://dx.doi.org/10.1016/s0024-3841(03)00044-5).

FRIEDERICI, Angela D.. The Brain Basis of Language Processing: From Structure to Function. **Physiological Reviews**, [s.l.], v. 91, n. 4, p.1357-1392, out. 2011. American Physiological Society. <http://dx.doi.org/10.1152/physrev.00006.2011>.

FRIEDERICI, Angela D.. **Language in our brain: the origins of a uniquely human capacity**. Cambridge, Massachusetts: The Mit Press, 2017.

FRITH, Uta; FRITH, Christopher D.. Development and neurophysiology of mentalizing. **Philosophical Transactions Of The Royal Society Of London. Series B: Biological Sciences**, [s.l.], v. 358, n. 1431, p.459-473, 4 fev. 2003. The Royal Society. <http://dx.doi.org/10.1098/rstb.2002.1218>.

GARCIA, Othon M.. **Comunicação em prosa moderna: aprenda a escrever, aprendendo a pensar**. 27. ed. Rio de Janeiro (RJ): Fgv, 2011. 548 p. Páginas 71-75.

GIL, David. Sign languages, creoles, and the development of predication. **Measuring Grammatical Complexity**, [s.l.], p.37-64, 30 out. 2014. Oxford University Press. <http://dx.doi.org/10.1093/acprof:oso/9780199685301.003.0003>.

GOBET, Fernand. *Entrenchment, Gestalt Formation, and Chunking*. In: SCHMID, Hans-jörg (Ed.). **Entrenchment and the Psychology of Language Learning**. [s.l.]: de Gruyter Mouton, 2017. p. 245-267.

GOLDFARB, Robert S.; RATNER, Jon. “Theory” and “Models”: Terminology Through the Looking Glass. **Economics Journal Watch**, Washington, v. 5, n. 1, p.91-108, jan. 2008. Disponível em: <<https://econjwatch.org/articles/-theory-and-models-terminology-through-the-looking-glass>> . Acesso em: 20 nov. 2019.

GOLDIN-MEADOW, Susan. **The resilience of language**: what gesture creation in deaf children can tell us about how all children learn language. New York: Taylor & Francis Group, 2005. 285 p.

GOLSTON, Chris; WIESE, Richard. The structure of the German root. **Phonology And Morphology Of The Germanic Languages**, [S.L.], p. 165-186, 31 dez. 1998. DE GRUYTER. <http://dx.doi.org/10.1515/9783110919769.165>.

GRODNER, Daniel; GIBSON, Edward. Consequences of the Serial Nature of Linguistic Input for Sentential Complexity. **Cognitive Science**, [s.l.], v. 29, n. 2, p.261-290, 4 mar. 2005. Wiley. http://dx.doi.org/10.1207/s15516709cog0000_7.

GROSZ, Barbara J.; SIDNER, Candace L.. ATTENTION, INTENTIONS, AND THE STRUCTURE OF DISCOURSE. **Computational Linguistics**, Cambridge, Ma, v. 12, n. 3, p.175-204, 1 jul. 1986.

GROSZ, Barbara J.; WEINSTEIN, Scott; JOSHI, Aravind K.. Centering: A Framework for Modeling the Local Coherence of Discourse. **Computational Linguistics**, Cambridge, Ma, v. 21, n. 2, p.203-225, jun. 1995.

HARRISON, Rebecca R.. A Structural Arrangement of Text to Facilitate Reading. **The Classical Journal**, Natchtoches, La, v. 102, n. 3, p. 291-303, 2007.

HARTMANN, Iren; HASPELMATH, Martin; TAYLOR, Bradley. **Valency Patterns Leipzig**. 2013. Disponível em: <http://valpal.info>. Acesso em: 19 jun. 2021.

HARTSUIKER, Robert J.. Are forward models enough to explain self-monitoring? Insights from patients and eye movements. **Behavioral And Brain Sciences**, [S.L.], v. 36, n. 4, p. 357-358, 24 jun. 2013. Cambridge University Press (CUP). <http://dx.doi.org/10.1017/s0140525x12002749>.

HASPELMATH, Martin. Arguments and Adjuncts as Language-Particular Syntactic Categories and as Comparative Concepts. **Linguistic Discovery**, [S.L.], v. 12, n. 2, p. 3-11, 2014. Dartmouth College Library Press. <http://dx.doi.org/10.1349/ps1.1537-0852.a.442>.

HASPELMATH, Martin; HARTMANN, Iren. Comparing verbal valences across languages. In: MALCHUKOV, Andrej; COMRIE, Bernard. **Introducing the Framework, and Case Studies from Africa and Eurasia**: volume 1 of comparative handbooks of linguistics. Berlin: Walter de Gruyter GmbH & Co, 2015. p. 41-70.

HERCULANO-HOUZEL, S.. The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. **Proceedings Of The National Academy Of Sciences**, [S.L.], v. 109, n. 1, p. 10661-10668, 20 jun. 2012. Proceedings of the National Academy of Sciences. <http://dx.doi.org/10.1073/pnas.1201895109>.

HILTON, Matt et al. Parallels in Processing Boundary Cues in Speech and Action. **Frontiers In Psychology**, [s.l.], v. 10, p.1-12, 16 jul. 2019. Frontiers Media SA. <http://dx.doi.org/10.3389/fpsyg.2019.01566>.

HOPPER, Paul J.; THOMPSON, Sandra A.. Transitivity in Grammar and Discourse. **Language**, [s.l.], v. 56, n. 2, p.251-299, jun. 1980. JSTOR. <http://dx.doi.org/10.2307/413757>.

HUETTIG, Falk; HARTSUIKER, Robert J.. Listening to yourself is like listening to others: external, but not internal, verbal self-monitoring is based on speech perception. **Language And Cognitive Processes**, [S.L.], v. 25, n. 3, p. 347-374, abr. 2010. Informa UK Limited. <http://dx.doi.org/10.1080/01690960903046926>.

HUETTIG, Falk; MCQUEEN, James M.. The tug of war between phonological, semantic and shape information in language-mediated visual search. **Journal Of Memory And Language**, [S.L.], v. 57, n. 4, p. 460-482, nov. 2007. Elsevier BV. <http://dx.doi.org/10.1016/j.jml.2007.02.001>.

IZRE'EL, Shlomo. Syntax, Prosody, Discourse and Information Structure: The case for Unipartite clauses — A View from Spoken Israeli Hebrew / Sintaxe, prosódia, discurso e estrutura informacional. **Revista de Estudos da Linguagem**, [s.l.], v. 26, n. 4, p.1675-1726, 11 out. 2018. Faculdade de Letras da UFMG. <http://dx.doi.org/10.17851/2237-2083.26.4.1675-1726>.

IZRE'EL, Shlomo *et al.* Segmenting speech. In: IZRE'EL, Shlomo *et al.* **In search of a basic unit of spoken language: a corpus-driven approach**. Amsterdam: John Benjamins Publishing Company, 2020. p. 1-34.

JACKENDOFF, Ray. In Defense of Theory. **Cognitive Science**, [s.l.], v. 41, p.185-212, 27 nov. 2015. Wiley. <http://dx.doi.org/10.1111/cogs.12324>.

KAHNEMAN, Daniel. **Thinking, fast and slow**. New York: Farrar, Straus And Giroux, 2011.

KAYNE, Richard. Why Isn't This a Complementizer? In: SVENONIUS, P. **Functional Structure from Top to Toe**. New York: Oxford University Press, 2010. Cap. 7. p. 188-231.

KETNER, Katherine Heidel. **Size restrictions in prosodic morphology**. 2006. 240 f. Tese (Doutorado) - Curso de Linguistics, University Of Cambridge, Cambridge, 2006.

KONDO, Tadahisa; MAZUKA, Reiko. Prosodic planning while reading aloud: On-line examination of Japanese sentences. **Journal Of Psycholinguistic Research**, [s.l.], v. 25, n. 2, p.357-381, mar. 1996. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/bf01708578>.

KRAHN, Albert Edward. **A new paradigm for punctuation**. 2014. 244 f. Tese (Doutorado) - Curso de Linguistics, University Of Wisconsin, Milwaukee, 2014.

KUHN, Melanie R. et al. Aligning Theory and Assessment of Reading Fluency: Automaticity, Prosody, and Definitions of Fluency. **Reading Research Quarterly**, [s.l.], v. 45, n. 2, p.230-251, 6 abr. 2010. Wiley. <http://dx.doi.org/10.1598/rrq.45.2.4>.

LACY, Paul de. Maximal Words and the Maori Passive. **Optimality Theory In Phonology**, [S.L.], p. 495-512, 2004. Blackwell Publishing Ltd. <http://dx.doi.org/10.1002/9780470756171.ch27>.

LAKOFF, George; JOHNSON, Mark. **Metaphors we live by**. 6. ed. Chicago: University Of Chicago Press, 2003.

LAKOFF, George. Explaining Embodied Cognition Results. **Topics In Cognitive Science**, [s.l.], v. 4, n. 4, p.773-785, 7 set. 2012. Wiley. <http://dx.doi.org/10.1111/j.1756-8765.2012.01222.x>.

LANDIS, J. Richard; KOCH, Gary G.. The Measurement of Observer Agreement for Categorical Data. **Biometrics**, [S.L.], v. 33, n. 1, p. 159, mar. 1977. JSTOR. <http://dx.doi.org/10.2307/2529310>.

LANGACKER, Ronald W. **Entrenchment in Cognitive Grammar**. In: SCHMID, Hans-jörg (Ed.). **Entrenchment and the Psychology of Language Learning**. [s.i.]: de Gruyter Mouton, 2017. p. 39-56.

LEHRMAN, Robert A.. **The Political Speechwriter's Companion: A guide for writers and speakers**. Washington (dc): Cq Press, 2010. 362 p.

LEITE, Camila Tavares. **A relação entre a compreensão e os aspectos prosódicos na leitura em voz alta de falantes do PE e do PB**. 2012. 214 f. Tese (Doutorado) - Curso de Linguística, Universidade Federal de Minas Gerais, Belo Horizonte (Mg), 2012. Disponível em: <https://repositorio.ufmg.br/bitstream/1843/LETR-8U7NQ9/1/1308d.pdf>. Acesso em: 10 out. 2020.

LEVELT, W. Monitoring and self-repair in speech. **Cognition**, [S.L.], v. 14, n. 1, p. 41-104, jul. 1983. Elsevier BV. [http://dx.doi.org/10.1016/0010-0277\(83\)90026-4](http://dx.doi.org/10.1016/0010-0277(83)90026-4).

LEVIN, Beth. Argument Structure. **Linguistics**, [S.L.], 28 maio 2013. Oxford University Press. Disponível em: <http://dx.doi.org/10.1093/obo/9780199772810-0099>. Acesso em: 22 maio 2021.

LIMA, Aveliny Mantovan; CONSTANTINI, Ana Carolina. Prosódia e fonoaudiologia: do fonoestilo ao transtorno da linguagem. **Prosódia da Fala**, [S.L.], p. 133-144, 20 jun. 2017. Editora Edgard Blucher. <http://dx.doi.org/10.5151/9788580392593-08>.

LOBATO, Lúcia Maria Pinheiro. **Sintaxe gerativa do português: da teoria padrão à teoria da regência e ligação**. Belo Horizonte, Mg: Vigília, 1986. 558 p.

LOGACEV, Pavel; VASISHTH, Shravan. A Multiple-Channel Model of Task-Dependent Ambiguity Resolution in Sentence Comprehension. **Cognitive Science**, [s.l.], v. 40, n. 2, p.266-298, 30 mar. 2015. Wiley. <http://dx.doi.org/10.1111/cogs.12228>.

LOGIE, Robert H.; BELLETIER, Clément; DOHERTY, Jason M. Integrating Theories of Working Memory. In: LOGIE, Robert; CAMOS, Valérie; COWAN, Nelson. **Working Memory: state of the science**. Oxford: Oxford University Press, 2021. Cap. 14. p. 389-430.

LOTEM, Arnon; KOLODNY, Oren; HALPERN, Joseph Y.; ONNIS, Luca; EDELMAN, Shimon. The bottleneck may be the solution, not the problem. **Behavioral And Brain Sciences**, [S.L.], v. 1, n. 39, p. 39-39, 2016. Cambridge University Press (CUP). <http://dx.doi.org/10.1017/s0140525x15000886>.

LYONS, John. **Language and linguistics: an introduction**. Cambridge, Uk: Cambridge University Press, 1981. 356 p.

LYONS, John. **Lingua(gem) e linguística: uma introdução**. Rio de Janeiro, Rio de Janeiro: Zahar Editores, 1982. 322 p. Tradução: Marilda Winkler Averbug e Clarisse Sieckenius de Souza.

MACDONALD, Maryellen C.. How language production shapes language form and comprehension. **Frontiers In Psychology**, [S.L.], v. 4, p. 1-16, 2013. Frontiers Media SA. <http://dx.doi.org/10.3389/fpsyg.2013.00226>.

MACLURE, Malcolm. The Case-Crossover Design: A Method for Studying Transient Effects on the Risk of Acute Events. **American Journal Of Epidemiology**, [s.l.], v. 133, n. 2, p.144-153, 15 jan. 1991. Oxford University Press (OUP). <http://dx.doi.org/10.1093/oxfordjournals.aje.a115853>.

MACLURE, M.; MITTLEMAN, And M. A.. Should We Use a Case-Crossover Design? **Annual Review Of Public Health**, [s.l.], v. 21, n. 1, p.193-221, maio 2000. Annual Reviews. <http://dx.doi.org/10.1146/annurev.publhealth.21.1.193>.

MAIA, Marcus; FINGER, Ingrid (Org.). **PROCESSAMENTO DA LINGUAGEM**. Pelotas, Rs: Educat - Editora da Ucpel, 2005. 535 p. (Investigações em Psicolinguística GT de Psicolinguística da ANPOLL).

MAIA, Marcus. PROCESSAMENTO DE FRASES E CONTEXTO DISCURSIVO. **Alfa: revista de linguística**, [s.i.], v. 63, n. 3, p.559-582, mar. 2019. Disponível em: <<https://periodicos.fclar.unesp.br/alfa/article/view/10627>> . Acesso em: 23 nov. 2019.

MÄNNEL, Claudia; FRIEDERICI, Angela D.. Neural correlates of prosodic boundary perception in German preschoolers: If pause is present, *pitch* can go. **Brain Research**, [s.l.], v. 1632, p.27-33, fev. 2016. Elsevier BV. <http://dx.doi.org/10.1016/j.brainres.2015.12.009>.

MAX PLANCK INSTITUTE FOR PSYCHOLINGUISTICS (The Netherlands) (Ed.). **ELAN**: professional tool for the creation of complex annotations on video and audio resources. Disponível em: <<https://tla.mpi.nl/tools/tla-tools/elan/>> . Acesso em: 03 dez. 2019.

MENON, Odete Pereira da Silva. O sistema pronominal do português do brasil. **Letras**, Curitiba, v. 1, n. 44, p. 91-106, jan. 1995.

MERLO, Paola; FERRER, Eva Esteve. The Notion of Argument in Prepositional Phrase Attachment. **Computational Linguistics**, [S.L.], v. 32, n. 3, p. 341-378, set. 2006. MIT Press - Journals. <http://dx.doi.org/10.1162/coli.2006.32.3.341>.

MERLO, Sandra. **DINÂMICA TEMPORAL DE PAUSAS E HESITAÇÕES NA FALA SEMI-ESPONTÂNEA**. 2012. 179 f. Tese (Doutorado) - Curso de Linguística, Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, Campinas, Sp, 2012. Disponível em: <http://repositorio.unicamp.br/bitstream/REPOSIP/271141/1/Merlo_Sandra_D.pdf> . Acesso em: 01 dez. 2019.

MILLER, George A.. The magical number seven, plus or minus two: some limits on our capacity for processing information.. **Psychological Review**, [S.L.], v. 63, n. 2, p. 81-97, 1956. American Psychological Association (APA). <http://dx.doi.org/10.1037/h0043158>.

MORE: Mecanismo online para referências, versão 2.0. Florianópolis: UFSC Rexlab, 2013. Disponível em: < <http://www.more.ufsc.br/> > . Acesso em: 02/12/2019.

MORFORD, Jill P.. Grammatical development in adolescent first-language learners. **Linguistics**, [S.L.], v. 41, n. 4, p. 681-721, 13 jan. 2003. Walter de Gruyter GmbH. <http://dx.doi.org/10.1515/ling.2003.022>.

MUSILYU, Oyedeji et al. **ELECTROPHYSIOLOGICAL CORRELATES OF PROSODIC BOUNDARIES AT DIFFERENT LEVELS IN BRAZILIAN PORTUGUESE**. Disponível em:

<https://www.researchgate.net/profile/Musiliyu_Oyedeji/publication/334947648_ELECTROPHYSI

OLOGICAL_CORRELATES_OF_PROSODIC_BOUNDARIES_AT_DIFFERENT_LEVELS_IN_BRAZILIAN_PORTUGUESE/links/5d45eea94585153e593aedc5/ELECTROPHYSIOLOGICAL-CORRELATES-OF-PROSODIC-BOUNDARIES-AT-DIFFERENT-LEVELS-IN-BRAZILIAN-PORTUGUESE.pdf> . Acesso em: 01 dez. 2019.

NIVRE, Joakin. Two Notions of Parsing. In: ARPPE, Antti et al. **Inquiries into Words, Constraints and Contexts: Festschrift for Kimmo Koskenniemi on his 60th Birthday**. Stanford, Ca: Cslipublications, 2005. Cap. 11. p. 111-120. (CSLI Studies in Computational Linguistics ONLINE). Disponível em: <<https://web.stanford.edu/group/cslipublications/cslipublications/koskenniemi-festschrift/kk-festschrift-all-2005.pdf>> . Acesso em: 22 nov. 2019.

O'CONNELL, Daniel C.; KOWAL, Sabine. Communicating with One Another: Toward a Psychology of Spontaneous Spoken Discourse. **Cognition And Language. A Series In Psycholinguistics**, [s.l.], p.1-266, 2008. Springer New York. <http://dx.doi.org/10.1007/978-0-387-77632-3>.

O'CONNELL, Daniel C.; KOWAL, Sabine. Fluency and Hesitation. **Communicating With One Another**, [s.l.], p.1-7, 2008. Springer New York. http://dx.doi.org/10.1007/978-0-387-77632-3_3.

OCTOPARSE. **Web Scraping Software**. Disponível em: <<https://www.octoparse.com/>> . Acesso em: 03 dez. 2019.

OVIATT, Sharon. Predicting spoken disfluencies during human-computer interaction. **Computer Speech & Language**, [S.L.], v. 9, n. 1, p. 19-35, jan. 1995. Elsevier BV. <http://dx.doi.org/10.1006/csla.1995.0002>.

PARTEE, Barbara. Topic, Focus and Quantification. **Semantics And Linguistic Theory**, [S.L.], v. 1, p. 159, 3 abr. 1991. Linguistic Society of America. <http://dx.doi.org/10.3765/salt.v1i0.2918>.

PATEL, Aniruddh D.. MUSICAL RHYTHM, LINGUISTIC RHYTHM, AND HUMAN EVOLUTION. **Music Perception: An Interdisciplinary Journal**, [s.l.], v. 24, n. 1, p.99-104, set. 2006. University of California Press.

PATEL, Aniruddh D. et al. Experimental Evidence for Synchronization to a Musical Beat in a Nonhuman Animal. **Current Biology**, [s.l.], v. 19, n. 10, p.827-830, maio 2009. Elsevier BV. <http://dx.doi.org/10.1016/j.cub.2009.03.038>.

PAYNE, Thomas E.. **Describing morphosyntax: a guide for field linguists**. Cambridge: Cambridge University Press, 1997

PAYNE, Thomas E.. **Exploring language structure: a student's guide**. Cambridge (UK): Cambridge University Press, 2006. 367 p.

PCX. **Plagiarism checker X**. Disponível em: <<https://plagiarismcheckerx.com/>> . Acesso em: 3 dez. 2019.

PESETSKY, David. **Zero syntax: experiencers and cascades**. [s.i.]: Massachusetts Institute Of Technology, 1995. (Current studies in linguistics).

PICKERING, Martin J.; GARROD, Simon. An integrated theory of language production and comprehension. **Behavioral And Brain Sciences**, [s.l.], v. 36, n. 4, p.329-347, 24 jun. 2013. Cambridge University Press (CUP). <http://dx.doi.org/10.1017/s0140525x12001495>.

PIMENTEL, Tiago; ROARK, Brian; COTTERELL, Ryan. Phonotactic Complexity and Its Trade-offs. **Transactions Of The Association For Computational Linguistics**, [S.L.], v. 8, p. 1-18, dez. 2020. MIT Press - Journals. http://dx.doi.org/10.1162/tacl_a_00296.

PLATÃO. **Fedro**. [ca 360 a.C.]. Tradução de Benjamin Jowett. Disponível em: <<http://classics.mit.edu/Plato/phaedrus.html>> . Acesso em: 17 nov. 2019.

RAMBOW, Owen. The Simple Truth about Dependency and Phrase Structure Representations: an opinion piece. **Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the ACL**, Los Angeles, Califórnia, v. 1, n. 1, p. 337-340, 1 jan. 2010.

RASO, Tommaso. Fala e escrita: meio, canal, consequências pragmáticas e linguísticas. **Domínios de Lingu@Gem**, [S.L.], v. 7, n. 2, p. 12-46, 2013. EDUFU - Editora da Universidade Federal de Uberlândia. <http://dx.doi.org/10.14393/dl>.

RASO, Tommaso; MITTMANN, Maryualê Malvessi; MENDES, Anna Carolina Oliveira. O papel da pausa na segmentação prosódica de corpora de fala. **Revista de Estudos da Linguagem**, [S.L.], v. 23, n. 3, p. 883-922, 22 dez. 2015. Faculdade de Letras da UFMG. <http://dx.doi.org/10.17851/2237-2083.23.3.883-922>.

RASO, Tommaso; TEIXEIRA, Bárbara; BARBOSA, Plínio. Modelling automatic detection of prosodic boundaries for brazilian portuguese spontaneous speech. **Journal Of Speech Sciences**, [S.L.], v. 9, p. 105-128, 9 set. 2020. Universidade Estadual de Campinas. <http://dx.doi.org/10.20396/joss.v9i00.14957>.

RIBEIRO, Antonio João Carvalho. A TEORIA DO GARDEN-PATH E EVIDÊNCIAS DO PORTUGUÊS DO BRASIL. **Revista Prolíngua**, [s.i.], v. 10, n. 1, p.93-105, jan. 2015.

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL <http://www.rstudio.com/>.

SARDINHA, Tony Berber; MOREIRA FILHO, José Lopes; ALAMBERT, Eli. Corpus **brasileiro**. Disponível em: <<http://corpusbrasileiro.pucsp.br/cb/Inicial.html>> . Acesso em: 03 dez. 2019.

SARMENTO, Leila Lauar. **Gramática em textos**. 3. ed. São Paulo: Moderna, 2012.

SCARPA, Ester. A estrutura prosódica das disfluências em português brasileiro. **Cadernos de Estudos Linguísticos**, Campinas, v. 1, n. 54, p. 25-40, jun. 2012.

SCHIEFER, Ana Maria; ARCURI, Cláudia Fassin. **Avaliação da Fluência da Fala**. In: MARCHESAN, Irene Queiroz; SILVA, Hilton Justino da; TOMÉ, Marileda Cattelan. Tratado de Especialidades em Fonoaudiologia. São Paulo: Guanabara Koogan, 2014

SCHMID, Hans-jörg. A Framework for Understanding Linguistic *Entrenchment* and Its Psychological Foundations. In: SCHMID, Hans-jörg (Ed.). **Entrenchment and the Psychology of Language Learning**. [s.i.]: de Gruyter Mouton, 2017. p. 9-37.

SILVA, Carlos. O Sintagma Determinante nas línguas românicas: emergência, mudança e estrutura da periferia esquerda do SN. **Letras**, [S.L.], n. 60, p. 201-222, 13 out. 2020. Universidad Federal de Santa Maria. <http://dx.doi.org/10.5902/2176148543106>.

SILVA, Luis Filipe Lima e. O estatuto da sintaxe na fala: considerações acerca da proposta da language into act theory / the status of syntax in speech. **Revista de Estudos da Linguagem**, [S.L.], v. 28, n. 1, p. 271-330, 2 jan. 2020. Faculdade de Letras da UFMG. <http://dx.doi.org/10.17851/2237-2083.28.1.271-330>.

SIMONS, Daniel J.; LEVIN, Daniel T.. Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, [s.i.], v. 5, n. 4, p.644-649, jan. 1998.

STAUB, Adrian; CLIFTON, Charles. Syntactic prediction in language comprehension: Evidence from either...or.. **Journal Of Experimental Psychology: Learning, Memory, and Cognition**, [s.l.], v. 32, n. 2, p.425-436, 2006. American Psychological Association (APA). <http://dx.doi.org/10.1037/0278-7393.32.2.425>.

STAUB, Adrian; RAYNER, Keith. Eye movements and on-line comprehension processes. **The Oxford Handbook Of Psycholinguistics**, [S.L.], p. 326-342, 2 ago. 2007. Oxford University Press. <http://dx.doi.org/10.1093/oxfordhb/9780198568971.013.0019>.

STEFANOWITSCH, Anatol; FLACH, Susanne. The Corpus-Based Perspective on *Entrenchment*. In: SCHMID, Hans-jörg (Ed.). **Entrenchment and the Psychology of Language Learning**. [s.i.]: de Gruyter Mouton, 2017. p. 101-127.

STEPHEN THORNTON. Stanford Encyclopedia Of Philosophy. **Karl Popper**. 2019. Disponível em: < <https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=popper> > . Acesso em: 19 nov. 2019.

ZOLTÁN GENDLER SZABÓ. Stanford Encyclopedia Of Philosophy. **Compositionality**. 2020. Disponível em: < <https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=compositionality> > . Acesso em: 30 mar. 2020.

THE UNIVERSITY OF CALIFORNIA MUSEUM OF PALEONTOLOGY. **Science at multiple levels**. Disponível em: <https://undsci.berkeley.edu/lessons/pdfs/how_science_works_p19.pdf> . Acesso em: 20 nov. 2019.

THOMAS UEBEL. Stanford Encyclopedia Of Philosophy. **Vienna Circle**. 2017. Disponível em: <<https://plato.stanford.edu/entries/vienna-circle/#RedFouTwoCriParReb>> . Acesso em: 19 nov. 2019.

TORPPA, Minna; EKLUND, Kenneth; SULKUNEN, Sari; NIEMI, Pekka; AHONEN, Timo. Why do boys and girls perform differently on PISA Reading in Finland? The effects of reading fluency, achievement behaviour, leisure reading and homework activity. **Journal Of Research In Reading**, [S.L.], v. 41, n. 1, p. 122-139, 27 jan. 2017. Wiley. <http://dx.doi.org/10.1111/1467-9817.12103>.

TRASK, Robert Lawrence. **A Dictionary of Grammatical Terms in Linguistics**. New York, Ny: Routledge, 1996.

TVERSKY, A.; KAHNEMAN, D.. Judgment under Uncertainty: Heuristics and Biases. **Science**, [s.l.], v. 185, n. 4157, p.1124-1131, 27 set. 1974. American Association for the Advancement of Science (AAAS). <http://dx.doi.org/10.1126/science.185.4157.1124>.

UENO, Mieko; POLINSKY, Maria. Does headedness affect processing? A new look at the VO–OV contrast. **Journal Of Linguistics**, [S.L.], v. 45, n. 3, p. 675-710, 30 set. 2009. Cambridge University Press (CUP). <http://dx.doi.org/10.1017/s0022226709990065>.

WATKINS, Michael J.. Models as toothbrushes. **Behavioral And Brain Sciences**, [S.L.], v. 7, n. 1, p. 86-86, mar. 1984. Cambridge University Press (CUP). <http://dx.doi.org/10.1017/s0140525x00026303>.

WINKWORTH, Alison L.; DAVIS, Pamela J.; ELLIS, Elizabeth; ADAMS, Roger D.. Variability and Consistency in Speech Breathing During Reading. **Journal Of Speech, Language, And Hearing Research**, [S.L.], v. 37, n. 3, p. 535-556, jun. 1994. American Speech Language Hearing Association. <http://dx.doi.org/10.1044/jshr.3703.535>.

WINKWORTH, Alison L.; DAVIS, Pamela J.; ADAMS, Roger D.; ELLIS, Elizabeth. Breathing Patterns During Spontaneous Speech. **Journal Of Speech, Language, And Hearing Research**, [S.L.], v. 38, n. 1, p. 124-144, fev. 1995. American Speech Language Hearing Association. <http://dx.doi.org/10.1044/jshr.3801.124>.

YAMASHITA, Hiroko; CHANG, Franklin. “Long before short” preference in the production of a head-final language. **Cognition**, [S.L.], v. 81, n. 2, p. B45-B55, set. 2001. Elsevier BV. [http://dx.doi.org/10.1016/s0010-0277\(01\)00121-4](http://dx.doi.org/10.1016/s0010-0277(01)00121-4).

ZWICKY, Arnold M.. Clitics and Particles. **Language**, [S.L.], v. 61, n. 2, p. 283, jun. 1985. JSTOR. <http://dx.doi.org/10.2307/414146>.