

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Faculdade de Letras
Programa de Pós-graduação em Estudos Linguísticos (PosLin)

Átila Augusto Soares Vital

**UMA PROPOSTA *CORPUS-DRIVEN* PARA A CLASSIFICAÇÃO DE
DISFLUÊNCIAS SEGUNDO A ESTRUTURA INFORMACIONAL DA FALA
ESPONTÂNEA**

Belo Horizonte
2025

Átila Augusto Soares Vital

**UMA PROPOSTA *CORPUS-DRIVEN* PARA A CLASSIFICAÇÃO DE
DISFLUÊNCIAS SEGUNDO A ESTRUTURA INFORMACIONAL DA FALA
ESPONTÂNEA**

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

Linha de pesquisa: Estudos Linguísticos Baseados em Corpora.

Orientador: Dr. Tommaso Raso
Coorientador: Dr. Bruno Rocha

Belo Horizonte
2025

V836p Vital, Átila Augusto Soares.
Uma proposta corpus-driven para a classificação de disfluências segundo a estrutura informacional da fala espontânea [manuscrito] / Átila Augusto Soares Vital. – 2026.
1 recurso online (124 f. : il., color.) : pdf.
Orientador: Tommaso Raso.
Coorientador: Bruno Neves Rati de Melo Rocha.
Área de concentração: Estudos Linguísticos.
Linha de Pesquisa: Estudos Linguísticos Baseados em Corpora.
Dissertação (mestrado) – Universidade Federal de Minas Gerais, Faculdade de Letras.
Apêndices: f. 108-118.
Bibliografia: f. 119-124.
Exigências do sistema: Adobe Acrobat Reader.

1. Atos de fala (Linguística) – Teses. 2. Linguística de corpus – Teses. 3. Distúrbios da fala – Teses. I. Raso, Tommaso. II. Rocha, Bruno Neves Rati de Melo. III. Universidade Federal de Minas Gerais. Faculdade de Letras. IV. Título.

CDD : 410



UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS

FOLHA DE APROVAÇÃO

Uma proposta corpus-driven para a classificação de disfluências segundo a estrutura informacional da fala espontânea

ÁTILA AUGUSTO SOARES VITAL

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Estudos Linguísticos, como requisito para obtenção do grau de Mestre em Estudos Linguísticos, área de concentração Linguística Teórica e Descritiva, linha de pesquisa Estudos Linguísticos Baseados em Corpora.

Aprovada em 26 de janeiro de 2026, pela banca constituída pelos membros:

Prof. Dr. Tommaso Raso - Orientador

UFMG

Prof. Dr. Bruno Neves Rati de Melo Rocha - Coorientador

UFMG

Prof. Dr. André Vinícius Lopes Coneglian

UFMG

Prof. Dr. Alessandro Panunzi

Università di Firenze

Belo Horizonte, 26 de janeiro de 2026.



Documento assinado eletronicamente por **Tommaso Raso, Professor do Magistério Superior**, em 29/01/2026, às 10:23, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Andre Vinicius Lopes Coneglian, Professor do Magistério Superior**, em 09/02/2026, às 10:11, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Bruno Neves Rati de Melo Rocha, Professor do Magistério Superior**, em 09/02/2026, às 10:53, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Alessandro Panunzi, Usuário Externo**, em 09/02/2026, às 18:17, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4912590** e o código CRC **5F2FC13A**.

AGRADECIMENTOS

Aos meus orientadores, Tommaso e Bruno, agradeço pelas discussões sempre atentas e pela confiança construída ao longo do tempo. Muitas coisas deram certo (neste trabalho e fora dele) por conta da dedicação e da paciência de vocês. Ao Bruno, agradeço pela sintonia desde o início da graduação, por ter me mostrado com clareza o que significa fazer um trabalho *corpus-driven*. Ao Tommaso, agradeço pelas oportunidades e pela incansável disposição.

Ao LEEL, que condiciona um ambiente pleno para a pesquisa acadêmica com projetos consolidados e estimulantes.

A Ísis Fiorilo e a Marília Mendes, companheiras de mestrado.

Ao João Rocha, à Gabriela Toledo, à Marianna Bicalho, à Alessandra Uchôas, ao Tim Marçal.

À Júlia Duarte, à Isabela Lee, ao Paulo Ferreira, ao Thiago Carvalho, ao Rafael Pires.

Ao Dalmo Buzato, por sugestões fundamentais.

Ao Fábio Granja, pelo apoio e pela amizade.

Ao Eduardo Moreira da Silva, pela parceria na Ciência Política enquanto a bolsa da Capes demorava a chegar.

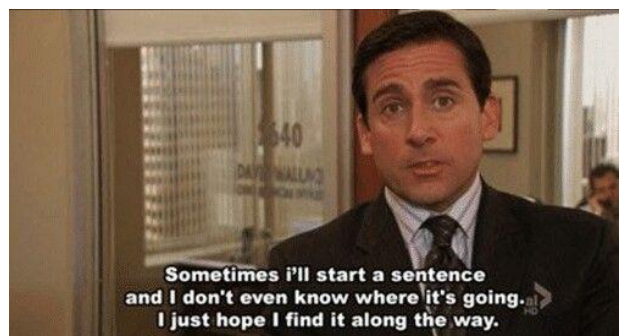
À Capes.

Ao grupo *Disfluencies in Spontaneous Speech* (DiSS), pela excelente discussão em Lisboa.

Aos membros da banca, pela disponibilidade e pelos excelentes comentários sobre o trabalho.

Aos amigos de vida.

À Suze, ao Átila, ao Thales e ao Toddy, por tudo.



(Michael Scott - The Office)

RESUMO

O objetivo deste trabalho é criar uma classificação de disfluências com base na estrutura informacional da fala espontânea. As disfluências são compreendidas como instâncias em que o falante suspende, temporariamente ou não, o fluxo de sua fala para reelaborá-la estrutural ou funcionalmente. Elas podem ser descritas etiologicamente - em função de suas causas na relação linguagem-pensamento (Chafe, 1980) - ou ontologicamente, através de variadas tentativas de classificação dos fenômenos (Schegloff, 2013; Kosmala, 2024). Neste trabalho, partimos da caracterização de diferentes tipos de disfluência e suas motivações em face da estrutura informacional, da maneira como ela é compreendida pela Language into Act Theory (LAcT) (Cresti, 2000; Moneglia & Raso, 2014). Com base em análises corpus-driven, foram definidas oito classes de disfluências e dez classes de motivações que podem se associar livremente, formando compostos disfluência-motivação. Todas as disfluências de uma amostra de 10 (dez) textos do minicorpus do C-ORAL-BRASIL I (Raso & Mello, 2012) foram etiquetadas manualmente com base nessa classificação. Foram identificados 944 compostos disfluência-motivação. A frequência relativa de cada uma das classes foi descrita em função da arquitetura do corpus e das unidades definidas pela L-AcT. Tanto os tipos de disfluência quanto as motivações se distribuem de maneiras distintas em face da estrutura informacional. As unidades informacionais que comportam com mais frequência os fenômenos de disfluência são o Parentético (46,7%), o Comentário Ligado (40,0%) e o Tópico (31,5%). A partir das medidas de frequência relativa, propomos uma organização dos fenômenos em dois níveis: o nível locutivo (como em um reparo morfossintático) e o nível informacional-ilocucionário (como em um reparo de unidade informacional ou de ilocução). As disfluências ocorridas no nível locutivo não têm implicação no nível informacional-ilocucionário; aquelas ocorridas no nível informacional-ilocucionário podem condicionar outras disfluências no nível locutivo (como pausas preenchidas e repetições). Essa proposta é sustentada pelos dados empíricos e evidencia que mesmo os fenômenos de fragmentação (as disfluências) atuam sobre a função e a forma na linguagem.

Palavras-chave: disfluências; fala espontânea; estrutura informacional; *Language into Act Theory*.

ABSTRACT

The goal of this work is to develop a classification of disfluencies based on the informational structure of spontaneous speech. Disfluencies are understood as instances in which the speaker suspends, either temporarily or not, the flow of speech in order to restructure it structurally or functionally. They may be described etiologically - in terms of their causes in the relationship between language and thought (Chafe, 1980) - or ontologically, through various attempts to classify the phenomena themselves (Schegloff, 2013; Kosmala, 2024). In this work, we begin from the characterization of different types of disfluency and their motivations in relation to informational structure, as understood within the Language into Act Theory (LAcT) (Cresti, 2000; Moneglia & Raso, 2014). Based on corpus-driven analyses, eight classes of disfluencies and ten classes of motivations were defined, which can combine freely to form disfluency-motivation compounds. All disfluencies in a sample of ten texts from the C-ORAL-BRASIL I minicorpus (Raso & Mello, 2012) were manually annotated according to this classification. A total of 944 disfluency-motivation compounds were identified. The relative frequency of each class was described with respect to the architecture of the corpus and the informational units defined by L-AcT. Both disfluency types and motivations exhibit distinct distributions in relation to informational structure. The informational units that most frequently host disfluency phenomena are the Parenthetical (46.7%), the Bound Comment (40.0%), and the Topic (31.5%). Based on relative-frequency measures, we propose an organization of the phenomena at two levels: the locutive level (as in morphosyntactic repairs) and the informational-illocutionary level (as in repairs of informational units or illocution). Disfluencies occurring at the locutive level have no implications for the informational-illocutionary level; those occurring at the informational-illocutionary level, however, may condition additional disfluencies at the locutive level (such as filled pauses and repetitions). This proposal is supported by empirical evidence and demonstrates that even phenomena of fragmentation (disfluencies) act upon both the function and the form of language.

Keywords: disfluencies; spontaneous speech; informational structure; Language into Act Theory.

LISTA DE FIGURAS

Figura 1. Frequência fundamental e forma de onda do exemplo 17.	41
Figura 2. Frequência fundamental e forma de onda do exemplo 18.	42
Figura 3. Frequência fundamental e forma de onda do exemplo 19.	43
Figura 4. Forma de onda e f_0 do exemplo 31.	71
Figura 5. Total de tipos de disfluências anotados.	86
Figura 6. Total de motivações anotadas.	87
Figura 7. Valor total e relativo de associações entre disfluências e motivações.	87
Figura 8. Frequência das motivações (%) associadas a cada uma das disfluências.	89
Figura 9. Disfluências encontradas nas tipologias monológica (azul), dialógica (laranja) e conversacional (verde).	91
Figura 10. Disfluências encontradas nos contextos privado (azul) e público (laranja).	92
Figura 11. Frequência relativa das unidades informacionais com disfluências.	94
Figura 12. Frequência relativa de unidades informacionais nas quais as disfluências ocorreram.	96
Figura 13. Frequência relativa de unidades informacionais nas quais as motivações ocorreram.	98
Figura 14. Frequência relativa de disfluências em cada uma das unidades ilocucionárias.	99
Figura 15. Frequência relativa de motivações de disfluências em cada uma das unidades ilocucionárias.	100

LISTA DE TABELAS

Tabela 1. Diferentes formas de se fazer referência às disfluências.	37
Tabela 2. Tipos de fenômenos anotados como disfluência no Corpus di Italiano Parlato.	53
Tabela 3. Tipos de fenômenos anotados como disfluências no C-ORAL-ROM.	54
Tabela 4. Tipos de fenômenos anotados como disfluências no C-ORAL-BRASIL I.	55
Tabela 5. Número de textos no minicorpus segundo as tipologias e os contextos.	59
Tabela 6. Gravações selecionadas do minicorpus para a anotação de disfluências.	60
Tabela 7. Unidades informacionais encontradas na amostra e suas frequências brutas.	93

LISTA DE SIGLAS

COM	Comentário
COB	Comentário Ligado
CMM	Comentário Múltiplo
TOP	Tópico
APC	Apêndice de Comentário
APT	Apêndice de Tópico
PAR	Parentético
ALL	Alocutivo
CNT	Conativo
DCT	Conector discursivo
EXP	Expressivo
AUX	Auxílio Dialógico
INP	Incipitário
INT	Introdutor Locutivo
TPL	Lista de Tópicos
PHA	Fático
UNC	Unidade informacional desconhecida
EMP	Unidade entonacional sem valor informacional
SCA	Escansão
TMT	Tomada de Tempo
TAG_r	Unidade informacional reportada
i-TAG	Unidade informacional interrompida
PAR-TAG	Unidade informacional realizada no nível parentético
f ₀	Frequência fundamental
SBC	<i>Santa Barbara Corpus of Spoken American English</i>
L-AcT	<i>Language into Act Theory</i>
DiSS	<i>Workshop on Disfluencies in Spontaneous Speech</i>
IPO	<i>Institute for Perception Research/Eindhoven</i>
DB-CoM	<i>Database for Corpora Multimidia</i>
LEEL	Laboratório de Estudos Empíricos e Experimentais da Linguagem
LABLITA	Laboratório de Italianística da Universidade de Florença
PoS	<i>Part of Speech</i>

LISTA DE SÍMBOLOS

*	Início de turno
ABC	Sigla do falante nos corpora da família C-ORAL
/	Quebra prosódica não-terminal
//	Quebra prosódica terminal
[/n]	Retracting no C-ORAL-BRASIL
+	Enunciado interrompido
< >	Trecho sobreposto
hhh	Ruído paralinguístico
xxx	Palavra ininteligível
yyyy	Trecho ininteligível
&	Início de palavra interrompida
&he	Pausa preenchida
[/]	Retracting com repetição total no C-ORAL-ROM
[//]	Retracting com repetição parcial no C-ORAL-ROM
[///]	Ambiguidade retracting/interrompido no C-ORAL-ROM
-?	Hesitação sem interrupção no Corpus di Italiano Parlato
+/.	Interrupção causada pelo interlocutor no Corpus di Italiano Parlato
...	Interrupção intencional (suspensão) no Corpus di Italiano Parlato
+/	Interrupção do programa informacional no Corpus di Italiano Parlato

SUMÁRIO

1. Introdução	16
1.1. A fala espontânea	17
1.2. As disfluências	18
1.2.1. Ambiguidades de anotações	19
1.3. A importância da estrutura informacional	21
1.4. O objetivo	22
1.4.1. Objetivos específicos	22
1.5. A organização deste trabalho	22
2. Fundamentação Teórica	23
2.1. As disfluências	23
2.1.1. A origem da disfluência na verbalização da experiência em Chafe	23
2.1.2. Operações de reparo em Schegloff (2013)	27
2.1.2.1. <i>Replacing</i>	28
2.1.2.2. <i>Inserting</i>	29
2.1.2.3. <i>Deleting</i>	29
2.1.2.4. <i>Searching</i>	30
2.1.2.5. <i>Parenthesizing</i>	30
2.1.2.6. <i>Aborting</i>	31
2.1.2.7. <i>Sequence-jumping</i>	31
2.1.2.8. <i>Recycling</i>	32
2.1.2.9. <i>Reformatting</i>	33
2.1.2.10. <i>Reordering</i>	34
2.1.3. A proposta de Kosmala (2024)	34
2.1.4. A discussão terminológica	37
2.2. Language into Act Theory	39
2.2.1. As unidades informacionais	44
2.2.1.1. As unidades textuais	44
2.2.1.1.1. Comentário (COM)	44
2.2.1.1.2. Tópico (TOP)	45
2.2.1.1.3. Apêndice de Comentário (APC)	45
2.2.1.1.4. Apêndice de Tópico (APT)	46
2.2.1.1.5. Parentético (PAR)	46
2.2.1.1.6. Introdutor Locutivo (INT)	47

2.2.1.1.7. Comentário Múltiplo (CMM).....	47
2.2.1.1.8. Comentário Ligado (COB) e a stanza	47
2.2.1.2. As unidades dialógicas	48
2.2.2. Sintaxe e macrossintaxe na L-AcT	49
2.2.3. As disfluências na L-AcT	50
2.2.3.1. Anotações no <i>Corpus di Italiano Parlato</i>	50
2.2.3.2. Anotações no C-ORAL-ROM.....	53
2.2.3.3. Anotações no C-ORAL-BRASIL.....	54
3. Metodologia	57
3.1. C-ORAL-BRASIL I.....	57
3.1.1. O minicorpus etiquetado informacionalmente	59
3.2. A anotação das disfluências.....	61
3.2.1. As classes	63
3.3. A quantificação	66
4. Resultados.....	68
4.1. A classificação <i>corpus-driven</i>	68
4.1.1. As categorias de disfluência	68
4.1.1.1. <i>Repair</i> (reparo)	68
4.1.1.2. <i>Scanning</i> (escansão)	71
4.1.1.3. <i>Repetition</i> (repetição).....	72
4.1.1.4. <i>Abandonment</i> (abandono).....	73
4.1.1.5. <i>Filled pause</i> (pausa preenchida).....	74
4.1.1.6. <i>Change of textual planning</i> (mudança de planejamento textual)	75
4.1.1.7. <i>Change of informational planning</i> (mudança de planejamento informacional).....	76
4.1.2. As categorias de motivação	78
4.1.2.1. <i>Prospection</i> (prospecção)	78
4.1.2.2. <i>Inadequacy</i> (inadequação)	79
4.1.2.3. <i>Articulation problem</i> (problema articulatorio).....	80
4.1.2.4. <i>Improvement</i> (melhoramento)	80
4.1.2.5. <i>Uncertainty</i> (incerteza).....	81
4.1.2.6. <i>Program alteration</i> (alteração do programa)	81
4.1.2.7. <i>External interrupt</i> (interrupção externa).....	82
4.1.2.8. <i>Immediate response</i> (resposta imediata).....	83
4.1.2.9. <i>Satisfaction</i> (contentamento/satisfação)	85
4.2. Caracterização estatística da amostra.....	85
4.3. As disfluências e suas motivações	88

4.4. As disfluências e a arquitetura da amostra	90
4.5. As disfluências e as unidades informacionais	92
4.6. Uma proposta <i>corpus-driven</i> para o tratamento das disfluências.....	100
5. Considerações finais	105
APÊNDICE A	107
APÊNDICE B	108
6. Referências.....	118

1. Introdução

Este trabalho tem como objetivo elaborar uma classificação de tipos de disfluências da fala e de suas potenciais motivações. Para isso, realizamos estudos em duas etapas, lançando mão de corpora de fala espontânea do português brasileiro. As classes de disfluências foram geradas com base na organização do padrão informacional, da forma como é descrito pela *Language into Act Theory* (Cresti, 2000; Moneglia & Raso, 2014). A partir de um minicorpus anotado, partimos para uma análise *corpus-based*, que visou quantificar os fenômenos de disfluência classificados e cruzar os dados quantitativos com anotações pragmáticas da estrutura informacional e do caráter ilocucionário das sequências linguísticas.

O estudo das disfluências tem sido realizado por diferentes tradições teóricas e metodológicas. Com a disponibilidade de importantes corpora de fala nas últimas décadas, surge também a necessidade de um sistema classificatório que dê conta da complexidade das interações face a face. Corpora de referência costumam ser acompanhados de materiais explicativos que descrevem as convenções utilizadas para transcrição, segmentação e alinhamento texto-som-gesto. Por conta de diferenças relevantes em suas arquiteturas e propósitos, a anotação de fenômenos de disfluência pode tomar contornos variados em cada um desses materiais. Embora exista um número cada vez maior de corpora de alto padrão cobrindo cada vez mais famílias linguísticas, o problema da incomparabilidade dificulta também as investigações interlinguísticas.

As diferenças com relação à granularidade e aos tipos de anotação de disfluências têm sido discutidas no âmbito da *Language into Act Theory* por ao menos duas tradições (italiana e brasileira). Com este trabalho, pretendemos fomentar a discussão e motivar uma classificação que se baseie na organização informacional, evidenciando a separabilidade entre os fenômenos em si e suas variadas motivações. Os corpora C-ORAL-BRASIL I (Raso & Mello, 2012), que documenta o português brasileiro falado informal, e o C-ORAL-ROM (Cresti & Moneglia, 2005), que documenta as quatro principais línguas românicas¹ faladas na Europa, foram utilizados como principais materiais para análise comparativa. Ambos também contam com seções anotadas informacionalmente segundo a *Language into Act Theory*.

¹ Italiano, francês, espanhol e português europeu.

Este texto conta com material complementar (áudio dos exemplos, tabelas de dados e corpus anotado) disponível em repositório no *Open Science Framework* (OSF)². No texto, os exemplos com arquivos de áudio estão nomeados com o ícone “🔊”. Além dos áudios, há os arquivos de alinhamento e as tabelas utilizadas no tratamento dos dados. O link do repositório se encontra no Anexo A.

1.1. A fala espontânea

A noção de fala espontânea utilizada neste trabalho diz respeito à fala cujo planejamento se dá durante a interação. Essa modalidade de comunicação representa a maior parte das nossas interações em contextos naturais, sendo caracterizada por diferentes disfasias e níveis de formalidade. Diferentemente da escrita, nas interações face a face, os interlocutores compartilham o *aqui e agora*, reduzindo ao mínimo a diferença entre os tempos de produção e o tempo de recepção. Estão também muito mais presentes nesse tipo de comunicação as pistas de natureza multimodal, como gestos e expressões faciais. Diferentemente da fala, a escrita é uma modalidade de uso da linguagem que permite reelaborações, revisão de escolhas vocabulares e correções *ad infinitum*, até que a versão final do texto seja utilizada ou esteja disponível aos leitores. O caráter imediato da fala diminui estrondosamente as possibilidades temporais para a realização de reelaborações e correções dos fenômenos que o falante julga serem imprecisos; embora isso esteja longe de ser um empecilho para a eficiência da comunicação.

A fala, quando executada espontaneamente, constitui a realização primordial da linguagem humana, muito mais antiga do que a escrita, que possui cerca de 5k anos. Tanto a estrutura conversacional (em turnos) quanto a construção dos enunciados (dentro dos turnos) parecem ter evoluído para maximizar a troca conversacional que ocorre quando os falantes compartilham o mesmo contexto extralinguístico (Levshina, 2022). Essas situações de comunicação geram uma série de fenômenos de interesse, como *backchannels*, sobreposições, *interactive repairs*, disfluências e etc. Algumas dessas instâncias, na realidade, possuem maior frequência do que as próprias palavras gramaticais em corpora de fala. Por essa razão, elas têm sido recentemente descritas como partículas fundamentais da interação (Dingemanse & Enfield, 2024), local restrito apenas aos itens gramaticais no caso da escrita. A escolha dos critérios para anotação desses elementos deve partir de uma descrição detalhada de suas distribuições e funções.

² <https://osf.io/c84mx/files/osfstorage>

1.2. As disfluências

Em *lato sensu*, as disfluências são compreendidas aqui como instâncias em que o falante suspende intencionalmente o fluxo da fala com o objetivo de realizar diferentes operações, como reparos na estrutura linguística, mudanças no planejamento e reformulações de ideias. Essa conceituação será expandida nos próximos capítulos, de modo a comportar tanto as particularidades formais dos fenômenos quanto seus objetivos.

São vários os termos utilizados na literatura para fazer referência a esses fenômenos; cada um deles privilegia certos aspectos e pontos de vista com relação à forma, frequência, grau de intencionalidade e potencial de convencionalização. Por conveniência, adotamos a nomenclatura mais frequente (de disfluência) e discutimos em detalhe na fundamentação teórica outras acepções, como hesitações (Chafe, 1980; Eklund, 2004), fenômenos de fragmentação (Cresti & Moneglia, 2005), a recente noção de inter-(dis)fluência (Kosmala, 2024), entre outros.

A título de uma primeira ilustração, o exemplo 1 representa o trecho de um diálogo informal selecionado do *Santa Barbara Corpus of Spoken American English* (Du Bois et al., 2000). Nele, a falante produz duas unidades entonacionais separadas na transcrição por uma quebra de linha. A primeira acaba sendo suspensa após a produção do primeiro segmento (‘d-’) de uma das palavras (provavelmente, *different*). Na sequência, se realiza a segunda unidade entonacional, que reelabora parte do conteúdo locutivo utilizado para expressar o mesmo pensamento de modo intensificado.

Exemplo 1 SBC001

LYNNE: that's a d- --

LYNNE: a whole different thing.

Com um exercício inferencial, é possível pensar que a primeira unidade entonacional tenha sido planejada como no exemplo 2, mas foi intencionalmente reelaborada na segunda unidade, incluindo um paradigma de função intensificadora (*whole*). Com isso, as interpretações semântica e morfosintática são realizadas desconsiderando os itens representados aqui como sombreados.

Exemplo 2 Exercício inferencial

LYNNE: that 's a different thing

LYNNE: that 's a whole different thing

Nessa perspectiva, a disfluência em questão teve como objetivo alterar o modo de apresentação da informação. Para isso, foram salvos o pronome demonstrativo e o verbo de ligação pronunciados na primeira unidade entonacional; foi acrescentado à segunda unidade um paradigma preenchido pelo intensificador “*whole*”, seguido da pronúncia completa do objeto sintático “*different thing*”.

Casos como esse são documentados extensivamente em corpora de fala e representam apenas um dentre os vários tipos de disfluências da fala. Além do reuso de estruturas linguísticas, acréscimo ou substituição de palavras e segmentos, outros elementos podem estar presentes ou mesmo constituírem disfluências isoladamente, como pausas preenchidas e repetições. Em todo caso, no entanto, são perceptíveis a suspensão ou o abandono da sequência linguística em construção. Embora muito frequentes, são raros os momentos em que esses fenômenos geram problemas de interpretação, o que sustenta a hipótese de que a co-construção da atividade comunicativa envolve o reconhecimento desses elementos e das intenções por parte do falante ao realizá-los. Do mesmo modo, potenciais imprecisões na veiculação da informação que não sejam suficientemente solucionadas em momentos de disfluências acabam sendo resolvidas através dos chamados *interactive repairs* (Dingemanse & Enfield, 2024).

1.2.1. Ambiguidades de anotações


Exemplos de fenômenos similares ao do exemplo 1 podem ser encontrados em corpora da família C-ORAL³, com anotações naturalmente diferentes daquelas utilizadas pelo corpus *Santa Barbara*. Os critérios de transcrição e as anotações do C-ORAL-BRASIL I documentam quatro tipos de disfluência: pausas preenchidas (&he), *retractings* ([/]), sequências interrompidas (+) e escansões (SCA)⁴. Além das pausas preenchidas, o C-ORAL-ROM enumera três tipos dos chamados fenômenos de fragmentação: as interrupções (+), os *retractings/false-starts* ([/] e [//]) e casos de ambiguidade entre *retractings* e sequências interrompidas ([///]). Cada um desses elementos será exemplificado no capítulo 2.

Com os trabalhos de discussão do grupo C-ORAL-BRASIL a respeito das disfluências, identificamos uma série de fenômenos que são rotulados com as etiquetas fornecidas pelos

³ A família C-ORAL é o conjunto de corpora produzidos segundo as diretrizes básicas desenvolvidas pelo LABLITA para a compilação do C-ORAL-ROM e aplicadas posteriormente para a compilação de outros corpora, como o C-ORAL-BRASIL.

⁴ As escansões são tradicionalmente entendidas como consequência da não correspondência biunívoca entre uma unidade informacional e uma unidade tonal (Cresti, 2000). Essa correspondência é um dos fundamentos da fala espontânea, e seu não seguimento pode ser analisado de diferentes maneiras, a depender da natureza da unidade informacional escansionada (se ilocucionária ou não). Esses fenômenos serão melhor introduzidos na fundamentação teórica.

corpora em questão, embora possuam propriedades muito diversas. Da mesma forma como uma anotação de *retracting* pode comportar um reparo gramatical (no exemplo 3, uma correção de concordância nominal), ela também pode comportar uma simples repetição de palavras funcionais (exemplo 4). Embora esses dois fenômenos sejam anotados da mesma forma (como *retractings*), eles não parecem ser tão similares, exceto pelo fato de envolverem itens funcionais. O oposto, isto é, o mesmo fenômeno sendo anotado de duas formas diferentes, ocorre quando há um reparo que envolve a repetição de palavras anteriores ao item reparado e aquele em que essa repetição não ocorre. Segundo as classificações do C-ORAL-ROM e do *Corpus di Italiano Parlato*, o primeiro seria classificado como *retracting* ([/] ou [//]), enquanto que o segundo se enquadraria como um caso de ambiguidade *retracting*-enunciado interrompido ([///]) (exemplo 5). Durante a análise dos dados, o que temos percebido é que essa repetição de palavras durante um reparo não parece ser necessária para a caracterização do próprio reparo. Na realidade, se trata de um fenômeno probabilístico, que apenas pode ocorrer quando ocorrem reparos. Uma das hipóteses aventadas nesta dissertação é a de que tais repetições ocorrem em função dos fenômenos de ilhas sintáticas causadas pelas unidades informacionais (cf. seção 4.6).

Exemplo 3 *bfamd131* (C-ORAL-BRASIL I) 

*LIQ: *cê viu como é que ficou lá os [//] as suculentas //*

Exemplo 4 *bfamcv28* (C-ORAL-BRASIL I) 

*ELI: *pois é / eu queria ir lá pra [//] pra olhar bolsa //*

Exemplo 5 *Corpus di Italiano Parlato*

*PNI: *simmeno &vo [///] spero / d' andare un po' in giro //*

A partir de uma análise *corpus-driven* de uma série de fenômenos de fragmentação, nossa classificação proporcionará a diferenciação entre as disfluências baseada na continuidade ou não das unidades informacionais. Até o momento, parece também ser possível estabelecer uma distinção entre fenômenos que dizem respeito ao nível estrutural (como reparos gramaticais, mudanças de planejamento e repetições) e aqueles que dizem respeito a funções pragmáticas (como reformulação da ilocução). Partindo disso, eliminamos uma série de problemas quanto à sobreposição de tipos de disfluências dentro de uma mesma classe. Na seção 4.6, propomos que os dois níveis sejam chamados locutivo e informacional-ilocucionário. Uma disfluência ocorrida no primeiro nível não tem efeito sobre o segundo; por outro lado, uma

disfluência ocorrida no nível informacional-ilocucionário pode gerar outra disfluência no nível locutivo. A discussão dos dados e o trabalho quantitativo em face das unidades informacionais parece sustentar ao menos a versão fraca dessa proposta.

Convém salientar que uma tentativa como essa já aparece em Cresti (2000), em que há uma primeira investida na descrição dos fenômenos de execução linguística e os fenômenos de fragmentação em termos (i) da unidade entonacional e (ii) do programa informacional. Por razões desconhecidas, esse trabalho já aparece com menos fôlego em Cresti & Moneglia (2005). O avanço da reflexão teórica no âmbito da L-AcT e o ganho de experiência para análise de corpora de fala constroem um terreno fértil para que se possa revisitar esses conceitos.

1.3. A importância da estrutura informacional

Em linhas gerais, a estrutura informacional (EI) diz respeito à organização dos blocos de informação ao longo do discurso (Raso, 2024). A nomenclatura *estrutura informacional* advém de Halliday (1967) e vem sendo referenciada também como empacotamento ou status informacional (Chafe, 1976). Em todo o caso, diz respeito a um nível de análise pragmática que se relaciona mais à maneira como a informação é transmitida por meio da estrutura linguística do que ao próprio conteúdo proposicional. Diferentes tradições de análise comportam descrições variadas sobre as possibilidades de empacotamento informacional.

Neste trabalho, lidamos com a noção de EI a partir da *Language into Act Theory*, que descreve uma série de unidades prosódicas que veiculam de forma convencional funções pragmáticas diversas. No ponto de vista da L-AcT, as unidades informacionais ocorrem em função de sequências que veiculam ilocuções (Austin, 1962). Em outras palavras, as unidades informacionais compõem o conteúdo semântico do enunciado e estabelecem as condições pragmáticas necessárias para que as ilocuções sejam interpretadas. A EI também se apresenta como nível linguístico determinante para a interpretação morfossintática; na fala, a prosódia (Chafe, 1994; Cresti, 2000) atua como veículo das funções informacionais independentemente do conteúdo semântico. A proposta da *Language into Act Theory* para a estrutura informacional será apresentada no capítulo 2.

Nosso ponto de vista tem como corolário a ideia de que a boa transmissão das funções informacionais das partes do enunciado é indispensável para a interpretação da ilocução. Embora a maior parte dos fenômenos de disfluência possuam alta frequência atestada em corpora de fala, postulamos que os falantes se esforçam para manter regularidades na execução de unidades informacionais. A classificação de disfluências com base em seus potenciais de interação com

as unidades informacionais fornece um panorama descritivo importante na determinação das relações entre a maneira como a informação é apresentada e sua boa construção linguística.

1.4. O objetivo

O principal objetivo deste trabalho é desenvolver uma classificação de disfluências ancorada na estrutura informacional da fala, do modo como ela é compreendida no âmbito da *Language into Act Theory*.

1.4.1. Objetivos específicos

Como objetivos específicos, cito:

1. Induzir uma tipologia de disfluências e uma tipologia de motivações a partir de uma análise *corpus-driven*;
2. Anotar as disfluências num conjunto de 10 textos provenientes do minicorpus do C-ORAL-BRASIL I segundo as classes de tipos e motivações desenvolvidas;
3. Verificar as unidades informacionais mais frequentes entre os diferentes tipos de disfluências e motivações;
4. Propor uma organização conceitual das disfluências da fala em face de propriedades locutivas e pragmáticas.

1.5. A organização deste trabalho

Este trabalho está organizado da seguinte forma. No capítulo 2, será apresentada a fundamentação teórica, que se divide em dois subcapítulos: o primeiro versa sobre a discussão do conceito de disfluência na literatura e contemporaneamente; o segundo apresenta os fundamentos da *Language into Act Theory*, a noção de padrão informacional e o tratamento dado pela teoria aos fenômenos de fragmentação. O terceiro capítulo é dedicado à metodologia, descrevendo os corpora utilizados, a construção da classificação de tipos e motivações de disfluências. O quarto capítulo apresenta os resultados da classificação e da análise baseada no corpus C-ORAL-BRASIL I. O quinto capítulo são as considerações finais.

Há também o apêndice 1, que diz respeito a um trabalho sobre a estrutura fonética e distribucional de pausas preenchidas para duas línguas românicas (português brasileiro e italiano). Este trabalho oferece uma consequência da compreensão de disfluências em face da estrutura informacional e da noção de convencionalização.

2. Fundamentação Teórica

Este capítulo tem como objetivo apresentar os fundamentos teóricos ao redor dos quais o trabalho foi realizado. Inicialmente, abordaremos uma proposta da natureza das disfluências da fala espontânea, discutindo sua etiologia (sua motivação) e sua ontologia (seus tipos). No segundo momento, apresentaremos a Language into Act Theory, teoria corpus driven que guia a anotação informacional utilizada na fase de anotação dos fenômenos. Após a apresentação das unidades informacionais, discutiremos os fenômenos de disfluência considerados pelas tradições italiana e brasileira.

2.1. As disfluências

Muitas vezes, as disfluências são descritas através de exemplificações (e não de conceituações) na literatura não especializada. Não é incomum que trabalhos que se utilizem de fala espontânea apresentem esses fenômenos da seguinte forma: “as disfluências são fenômenos de fragmentação que ocorrem no discurso, como eh..., alongamentos, reparos e autocorreções”. O problema de abordá-las dessa forma é precisamente a falta de uma conceitualização explícita, capaz de englobar cada um dos exemplos citados e ainda os que faltaram. A conceitualização depende naturalmente dos objetivos do estudo e do ponto de vista teórico. Aqui, optamos por apresentar um pequeno recorte partindo do modelo de verbalização da experiência em Chafe (1976; 1980; 1994) para a elucidação da etiologia das disfluências. Na sequência, serão apresentadas duas classificações, que versam sobre a ontologia dos fenômenos de disfluência: a classificação de Schegloff (2013) e a de Kosmala (2024). Para uma visão global desses elementos e de suas relações com a filosofia da linguagem, psicolinguística, psicose, gagueira, prosódia e muitas outras coisas, Eklund (2004) continua sendo uma fonte inspiradora, junto com todos os *proceedings* publicados desde 1999 pelo *Workshop on Disfluency in Spontaneous Speech* (DiSS).

2.1.1. A origem da disfluência na verbalização da experiência em Chafe

A ideia de W. Chafe por trás da verbalização da experiência advém de estudos da década de 1970 que também culminam nas chamadas *Pear Stories* (Chafe, 1980), realizadas na *University of California* (Berkeley). As *Pear Stories* são registros de fala semi-espontânea elicitadas a partir de um curta metragem⁵ mudo produzido para esse fim. Após assistirem ao filme, os falantes eram instruídos a narrar livremente os acontecimentos que tinham visto da

⁵ <https://www.youtube.com/watch?v=bRNSTxTpG7U&t=9s>

maneira como se lembravam. À época, o trabalho tinha como objetivo investigar o fluxo e os focos de consciência através de unidades entonacionais no âmbito de uma *thought-based linguistics*. A verbalização da experiência é o modelo desenvolvido por Chafe (2002; 2005, e anteriores) com o objetivo de fundamentar o surgimento da gramática segundo as relações de ajuste entre o pensamento e a linguagem.

Chafe (2002) parte da ideia de que os pensamentos (*thoughts*) veiculados no fluxo da fala são a força-motriz da linguagem, e não a sua estrutura. As ideias (*ideas*) dizem respeito aos diferentes tipos de experiências que um falante pode ter no mundo. Elas são de três tipos: ideias de eventos (o que acontece), ideias de estados (o que é ou está) e ideias de participantes (que experienciam eventos e estados). O pensamento possui algumas propriedades que encontram seus correlatos na fala espontânea (Chafe, 2005):

1. O pensamento é dinâmico.
2. O pensamento é segmentado em focos de consciência, delimitados na fala pelas unidades entonacionais.
3. O pensamento é um construto de ideias sobre eventos, estados e seus participantes.
4. As ideias são orientadas em várias dimensões, incluindo tempo, espaço, epistemologia, emoções, interações e contexto.

O pensamento, pela sua natureza, não pode ser expresso diretamente em sons da linguagem. Para isso, são postuladas cinco operações de ajuste responsáveis por fazer com que o conteúdo do pensamento se conforme às limitações semânticas e morfossintáticas impostas pela estrutura das línguas. Essas operações são:

1. Seleção
2. Categorização
3. Orientação
4. Combinação
5. Linearização

A seleção consiste no processo de recortar parte(s) do pensamento, considerando que ele é mais rico e mais complexo do que qualquer coisa que possa ser expressa pela linguagem. Uma situação prática, em que um falante foi a uma concessionária, fez *test-drives* em vários veículos e comprou um deles, pode ser expressa ao interlocutor como “eu comprei um carro econômico”. Nesse caso, toda a complexidade do pensamento associado às experiências vividas pelo falante foi condensada, e apenas algumas partes foram efetivamente selecionadas para comporem esse pequeno enunciado.

O processo de categorização faz com que um determinado pensamento seja expresso de maneira focalizada em termos de categorias definidas. No mesmo exemplo da concessionária, o falante seleciona e categoriza a propriedade de consumo de combustível, expressando que o carro comprado por ele é econômico. Algumas ideias são facilmente categorizáveis em estruturas prototípicas, como “comprar”, “carro” e “economia”. Outras impõem alguma dificuldade pela indisponibilidade da língua em comportar a unidade que se deseja representar. Esses casos parecem condicionar fenômenos de hesitação (Chafe, 2005).

O terceiro processo é chamado de orientação. As ideias que o falante possui acerca de determinada experiência são indexadas ao tempo, ao espaço, epistemologicamente, com relação a outras ideias e em face da interação. Quando ajustados à estrutura linguística, os pensamentos passam a ser orientados segundo essas instâncias. As línguas podem ter mecanismos variados para orientação dos pensamentos. Marcas temporais e espaciais orientam com relação ao tempo e ao espaço. Evidenciais orientam o pensamento epistemologicamente, indicando os sentidos envolvidos na aquisição da experiência (Chafe, 1986). A definitude orienta a identificabilidade ou não identificabilidade dos referentes. A prosódia, segundo Chafe (2002), faria parte de um mecanismo para a orientação do status informacional, além de ser responsável por acrescentar *evaluative orientations*⁶.

Os últimos dois processos de ajuste envolvidos na relação pensamento-linguagem são combinação e linearização. A primeira diz respeito às diferentes possibilidades de construções disponíveis nas línguas para combinar as ideias selecionadas, categorizadas e orientadas. A organização das funções sintáticas sujeito-verbo-objeto (e variações) e a queda de *pitch* ao final dos enunciados são correlatos da combinação e delimitam a veiculação de um pensamento completo. A linearização⁷ diz respeito à organização temporal dos sons da fala durante o processo de verbalização. Assim como a dependência do eixo temporal faz com que os segmentos sejam pronunciados e planejados em sequência, os pensamentos (formados por sequências de focos de consciência) passam pelo mesmo processo. No entanto, os focos de consciência em si, argumenta Chafe (2002), não possuem estrutura temporal interna. Ao contrário da maneira como se narra um acontecimento - linearmente -, o foco de consciência é apreendido como um todo, formado pelo evento e seus participantes.

⁶ É difícil afirmar com precisão o que Chafe quis dizer com *evaluative orientations*, mas tendo a interpretar como algo próximo a valores atitudinais.

⁷ Mais tarde, em Chafe (2005), a linearização é retirada dos processos de ajuste pensamento-linguagem. As razões desse movimento não ficam esclarecidas.

Todos os cinco processos envolvidos na verbalização da experiência em Chafe (2002) - com mais força na combinação e na linearização - parecem ser consequências da dicotomia saussuriana entre os eixos sintagmático e paradigmático. O resultado desses processos de ajuste é o que Chafe (2005) chama de estrutura semântica, isto é, o conjunto de elementos disponíveis numa dada língua para expressar partes do pensamento. Não é de se espantar que essas propriedades da linguagem impactem em maior ou menor grau também a interpretação dos fenômenos de disfluência.

A primeira tentativa por parte de Chafe de explicar detidamente a razão de ser das disfluências ocorre em Chafe (1980), *Some reasons for hesitating*. Esse texto surge mais de vinte anos antes da elaboração dos processos de verbalização da experiência, em Chafe (2002) e Chafe (2005). Por esse motivo, talvez, não sejam feitas referências explícitas aos cinco processos apresentados acima. No texto de 1980, os fenômenos são referenciados tanto como hesitações quanto como pausas⁸. Embora esses nomes ofereçam maior espaço às disfluências que envolvem a supressão do sinal acústico (as pausas), é ampla a gama de fenômenos considerados (pausas silenciosas e preenchidas, *false-starts*, *afterthoughts*, repetições e etc.). A preocupação inicial não é de estabelecer uma classificação de tipos, mas compreender o surgimento dessas categorias na verbalização (sua etiologia). A principal unidade de análise utilizada são as unidades entonacionais, que dizem respeito aos focos de consciência e têm a função de compartimentar a experiência numa sequência de ações discretas. Elas tendem a comportar até uma sentença completa (com exceções), além de exibirem variados contornos entonacionais. Os focos podem ser agrupados em *focus clusters*, que são estruturas mais complexas sintática e prosodicamente, observadas em diferentes contextos, como quando o falante descreve as estruturas espaciais de uma cena como se estivesse percorrendo com os olhos várias partes da imagem⁹.

As hesitações ocorreriam por conta de dois tipos de processos. Um deles se relaciona à procura do falante pelo próximo foco de consciência a ser verbalizado. O outro se relaciona à procura do falante pela melhor maneira de verbalizar o foco de consciência já selecionado. Em outras palavras, duas macro-motivações são arroladas para a realização de disfluências: a

⁸ O termo *pausologia* (*pausology*) é curiosamente utilizado para fazer referência ao estudo desses elementos, abrangendo não apenas as disfluências que envolvem pausas ou hesitações, mas a ampla gama de fenômenos que cai sob o conceito de disfluência.

⁹ É curioso notar como a ideia de *focus cluster* se aproxima em alguma medida da noção de stanzas ou mesmo de padrões informacionais no âmbito da *Language into Act Theory*. Os pontos de partida são bastante diferentes, mas a noção de que estruturas complexas organizadas em unidades entonacionais são geradas com variadas finalidades durante a fala funciona como ponto de aproximação. Na L-Act, as unidades informacionais orbitam unidades ilocucionárias, enquanto que, na proposta de Chafe, o *focus cluster* é gerado ao redor do foco de consciência.

procura *do que* dizer e a escolha de *como* dizer. Um interessante resultado quantitativo também realizado para narrativas espontâneas eliciadas por vídeos é o de que o maior número de hesitações está presente na mudança entre *focus clusters*. De forma inversa, um menor número de hesitações se encontra entre unidades entonacionais que pertencem ao mesmo cluster (Chafe, 1980).

Em suma, a proposta de W. Chafe para a etiologia das disfluências ocorre em função de dificuldades nos processos de ajuste entre o nível do pensamento e aquele da linguagem. Interessado em verbalizar a experiência, o falante usa variados mecanismos de reformulação na busca de acomodar da melhor forma possível o foco de consciência aos processos de seleção, categorização, orientação, combinação e linearização disponíveis para sua língua:

In this kind of investigation “ungrammatical”, “disfluent” speech constitutes the primary data, and hesitation phenomena are welcome as overt, measurable indications of processing activity which requires a certain amount of time. They provide good evidence that speaking is not a matter of regurgitating material already stored in the mind in linguistic form, but that it is a creative act, relating two media, thought and language, which are not isomorphic but require adjustments and readjustments to each other. A speaker does not follow a clear, well traveled path, but must find his way through territory not traversed before, where pauses, changes of direction, and retracing of steps are quite to be expected. The fundamental reason for hesitating is that speech production is an act of creation. (Chafe, 1980, p. 170).

O estudo das disfluências não se beneficia apenas por interesses classificatórios dos fenômenos em análise. As explicações etiológicas oferecem panoramas que estabelecem relações com outros mecanismos cognitivos, como é o caso da verbalização da experiência. Explicações de outras naturezas também estão disponíveis, principalmente no contexto psicolinguístico (a mais clássica é a de Levelt, 1983).

2.1.2. Operações de reparo em Schegloff (2013)

Muitas das classificações de disfluências disponíveis são desenvolvidas em trabalhos de Análise da Conversa. Essas propostas tomam o turno conversacional - e construtos derivados da noção de turno - como unidade de referência, de modo que os tipos de disfluências são categorizados em face do domínio do falante com relação à demanda de fala. Schegloff (2013) propõe dez classes de reparos realizados pelo próprio falante (*self-repairs*) - em oposição aos *other-initiated repairs* -, sem que haja troca de turnos. Os *self-repairs* são de especial interesse para esta dissertação, haja vista que um dos fundamentos da nossa proposta é categorizar as disfluências do ponto de vista das intenções do falante.

Os dez tipos¹⁰ são os seguintes: *replacing*, *inserting*, *deleting*, *searching*, *parenthesizing*, *aborting*, *sequence-jumping*, *recycling*, *reformatting* e *reordering*. Cada um deles será apresentado na sequência. Há ainda uma separação entre reparos de primeira ordem e reparos de segunda ordem. Os reparos de primeira ordem são aqueles que explicam com razoabilidade as operações estruturais executadas pelo falante ao longo da disfluência. A noção de reparo de segunda ordem, por outro lado, está relacionada ao papel dessas operações na dinâmica interacional. A princípio, todos os dez tipos apresentados podem ser interpretados como de primeira ou como de segunda ordem, a depender da maneira como são executados e da função que estabelecem na conversação, respectivamente. Em nossa apresentação, os critérios de transcrição dos exemplos foram simplificados, e apenas os trechos de interesse foram reproduzidos aqui.

2.1.2.1. *Replacing*

O *replacing* ocorre quando o falante, num mesmo turno conversacional, substitui uma sequência inteira ou parcialmente articulada por um ou outros elementos. Um dos exemplos de Schegloff (2013) é reproduzido na sequência.

Exemplo 6 (Schegloff, 2013)

*Bee: was I sid no I sid but we're supposetuh know what it
is for **Weh-**. hh yihknow for tihday's class 'n,*

Em 6, o falante substitui o trecho que provavelmente seria articulado como “*We[dnesday's] class*” por “*today's class*”. As operações de *replacing* podem envolver itens de mesma natureza gramatical ou não. Da mesma forma como se pode realizar um *replacement* de uma única palavra por outra, é possível também que esta mesma operação substitua uma palavra por um sintagma inteiro ou vice-versa. Além disso, podem haver repetições de trechos ditos anteriormente, como foi o caso de *for*, embora esta característica não seja necessária. Schegloff (2013) salienta o fato de que é comum que os trechos candidatos à operação de *replacing* sejam prematuramente terminados, isto é, marcados por algum tipo de interrupção da fala. Não fica claro em que medida esses tipos de interrupções geram, por exemplo, novas unidades entonacionais. Essa é uma das propriedades apontadas por Cresti (2000) e também por Cresti & Moneglia (2005) para a marcação de *retractings*.

¹⁰ Por se tratarem de termos bem estabelecidos na literatura da Análise da Conversa, optei por não traduzi-los.

2.1.2.2. *Inserting*

A operação de *inserting* ocorre quando o falante insere estruturas dentro de uma sequência que já havia sido iniciada ao longo do turno. Nesses casos, não há necessariamente uma correção a ser realizada, mas sim um melhoramento na feitura do turno. Por esse motivo, Schegloff (2013) aponta que esse tipo de disfluência não é bem representada pela categoria intitulada de reparo - que evoca a ocorrência de erros na estrutura do enunciado. A nomenclatura “*alteration*” talvez seja mais apropriada. O exemplo 7 representa a operação de *inserting*, em que a palavra *exactly* é inserida no turno após o início do primeiro segmento (*f*) de *funky*.

Exemplo 7 (Schegloff, 2013)

*Sta: Na- uh f not f **exactly** funky but*

Nesses casos, pode haver também a repetição de estruturas já pronunciadas antes da inserção ocorrer. Em termos conceituais, a diferenciação básica entre as *replacings* e as *insertions* reside justamente no fato de que as primeiras corrigem uma estrutura compreendida como problemática, enquanto que as segundas apenas melhoram a estrutura do turno. As diferenças, no entanto, não parecem de todo claras, e talvez sejam melhor compreendidas do ponto de vista da intencionalidade do falante.

2.1.2.3. *Deleting*

A operação de *deleting* ocorre quando o falante deleta um ou mais elementos já articulados no turno. Segundo Schegloff (2013), ela ocorre com muito menos frequência do que todos os outros tipos de reparo. O exemplo 8 mostra um caso em que a palavra semi-articulada *als[o]* foi deletada da estrutura.

Exemplo 8 (Schegloff, 2013)

*Bee: tuh go entry there. Because I **als-** I tried Barnes*

Nesse caso, tanto a estrutura *I also tried Barnes* quanto *I tried Barnes* parecem candidatas possíveis para a pronúncia. Embora isto não seja comentado explicitamente na lista de Schegloff (2013), é possível pensar que a deleção exija uma inferência maior quanto às palavras subsequentes ao termo deletado. Para que 8 seja classificado dessa forma, é necessário inferir que a estrutura previamente planejada fosse mesmo *I also tried Barnes*, ao contrário, por exemplo, de *I also thought about Barnes*. Este é o passo inferencial que parece ser necessário para a diferenciação entre *deleting* e *replacing*.

2.1.2.4. *Searching*

Segundo Schegloff (2013), o processo de *searching* diz respeito ao momento em que o falante está à procura de uma expressão linguística. Há diferentes subclassificações, como *precise searching*, *delicate searching* (Lerner, 2004) e *resumption searching* (Schegloff, 2013). O *precise searching* ocorre quando o falante procura por uma expressão linguística particular (geralmente um nome), enquanto que os *delicate searchings* ocorrem quando as expressões linguísticas em questão não são tão precisas. *Resumption searching* ocorre quando a busca se dá de forma concomitante a outros fenômenos, como interrupções e *side-trackings*. No exemplo 9, ocorre um caso de *precise searching*, em que o falante gasta uma parte do turno na procura do nome do teatro sobre o qual deseja falar.

Exemplo 9 (Schegloff, 2013)

Stn: *Bullocks? ya mean that one **right u:m tch!** (.)
right by thee u:m (.) was it Plaza? theatre::*

É importante notar que esse tipo de fenômeno, embora classificado como um tipo de reparo nesta tipologia, parece se comportar melhor como uma motivação para a realização de algumas disfluências. Como em 9, não são raros os casos em que o falante está indeciso a respeito da próxima palavra (e a está procurando) e, portanto, realiza uma pausa preenchida.

2.1.2.5. *Parenthesizing*

A operação chamada *parenthesizing* é descrita como um tipo de reparo, mas também ocorre com variadas funções. Através dessas estruturas, é possível fazer inserções de conteúdo no interior do turno, como em 10.

Exemplo 10 (Schegloff, 2013)

Sh1: no- well that was initially and then I'm like no:
I'll just go an then uhm you know thi- this tow
bandit (.) thing that I have, that were doing,
he wants me I- >**I don't know if I tol' you this,**<
he wants us to come out to his house and do:, like

De todas as operações de reparo descritas por Schegloff (2013) esta é uma das que mais se diverge, em função e forma, dos tipos de disfluência que serão propostas nesta dissertação. Isso ocorre por conta do fato de que os *parenthesizings* parecem ser classificados como tipos

de reparos pelo simples fato de inserirem conteúdo no fluxo do turno. Isso, no entanto, não exige claras mudanças de elaboração, marcas de imprecisão ou outros fenômenos típicos em casos de disfluência. A inserção de conteúdo locutivo através de estruturas parentéticas tem sido discutida inclusive no âmbito da L-AcT como uma importante propriedade pragmática (Rocha, 2025).

2.1.2.6. *Aborting*

A operação de *aborting* ocorre quando um turno (ou parte dele) é abandonado pelo falante. Schegloff (2013) faz uma separação de dois tipos. O primeiro se dá quando o falante o abandona na tentativa de reelaborá-lo, apresentando na sequência outra tentativa de construir o trecho. O segundo tipo ocorre quando não há tentativa subsequente de apresentar a informação. Nesse último caso, o turno é abandonado por completo. O exemplo 11 mostra um caso de *aborting* parcial, com tentativa de reelaboração.

Exemplo 11 (Schegloff, 2013)

Shr: Who's Debbie.

*Mrk: She's just that girl that uh,
hh I met her through uh m*

Ao ser questionado por Shr a respeito de uma terceira pessoa (Debbie), Mrk realiza a primeira tentativa para fornecer uma explicação. Na realização do trecho, o falante muda seu planejamento, realizando o *aborting*, e oferecendo a explicação de outra maneira.

2.1.2.7. *Sequence-jumping*

A operação de *sequence-jumping* envolve uma interrupção do turno em construção para uma mudança brusca de tópico. Isso pode ser feito em presença ou ausência de pausas e quebras. Diferentemente do *aborting*, ao realizar o *sequence-jumping*, o falante não tenta reelaborar o turno que estava sendo realizado. O exemplo 12 ilustra bem essa situação.

Exemplo 12 (Schegloff, 2013)

*Fre: You know what we're gonna in fact I'm- she I
haven't seen her since I spoke to you but I'm going to
talk to=what are you making?*

Kat: It's a blanket.

Motivadas pelas descrições da Análise da Conversa, tanto o fenômeno de *sequence-jumping* quanto os outros tipos de reparo são definidos majoritariamente com relação ao conteúdo informacional adjacentes à interrupção. Casos como o do exemplo 12, especificamente, ocorrem com maior frequência em contextos nos quais os falantes estão engajados em diferentes ações enquanto interagem (alta variação diafásica), além das ações realizadas estritamente através da linguagem. Desse ponto de vista, o salto informacional realizado por Fre pode ser também compreendido como uma rápida mudança acional frente aos estímulos do interlocutor. Esse tipo de explicação retira o peso do conteúdo semântico e o transfere às funções pragmáticas da fala espontânea.

2.1.2.8. *Recycling*

A operação de *recycling* ocorre por variados motivos e produz diferentes consequências, segundo Schegloff (2013). Ela consiste na repetição de trechos já realizados para a construção do turno e pode ou não ser acompanhada de outros fenômenos de reparo, como *insertings*. Nesses casos, os elementos envolvidos no *recycling* não constituem um tipo de reparo em si, mas funcionam como um recurso que pode ser utilizado pelo falante quando ocorrem disfluências. Reproduzimos abaixo o exemplo 6 (renumerado 13), utilizado na explicação de *replacing*.

Exemplo 13 (Schegloff, 2013)

*Bee: was I sid no I sid but we're supposetuh know what it
is for Weh-. hh yihknow for today's [class. 'n,*

Em 6, o *replacing* de *We[nesday]* por *today's* vem acompanhado pelo *recycling* da preposição *for*, que é repetida após a interrupção. Na classificação da disfluência que ocorre nesse caso, devemos pensar no *replacing*, enquanto o *recycling* atua como um recurso secundário para a boa formação do enunciado. O *recycling* também pode ocorrer isoladamente, sem envolver outros tipos de reparo. O exemplo 14 ilustra esse caso, em que o pronome *they* é repetido em sequência.

Exemplo 14 (Schegloff, 2013)

*Bee: No, I don't know I guess spring ter they don't
order- you know many book as they-they- really are
suppose to.*

Um questionamento relevante pode ser o motivo de não se fazer uma separação entre *recycling* e repetição. O primeiro não atua como um tipo de disfluência em si, mas como um recurso de retração da fala usado em contextos de disfluências diversas. O segundo, por sua vez, representa uma disfluência específica, marcando casos em que a repetição dos itens linguísticos ocorre por completo, sem alterações no conteúdo ou na estrutura gramatical. Distinções desse tipo são observadas em outras classificações, nas quais o tipo *repetition* já se encontra bem estabelecido.

2.1.2.9. *Reformatting*

A operação de *reformatting* é inicialmente apresentada por Schegloff (2013) como de natureza gramatical. Ela ocorre quando há algum tipo de reformulação no âmbito estrutural da gramática de parte do turno conversacional. O exemplo 15 ilustra um desses casos, em que Vir parece ter começado a produzir um enunciado declarativo, como em *Beth got to work [...]* ou ainda *Beth didn't have to wait [...]*, mas que acabou sendo alterado para uma forma interrogativa (*Didn't*) *Beth get to work before she was sixteen?*.

Exemplo 15 (Schegloff, 2013)

*Vir: Well **Beth didn't** Beth get to work before she was sixteen?*

Embora esse tipo de operação tenha sido descrito primeiramente como de natureza gramatical, é também possível pensar que a mudança estrutural ocorre motivada por um rearranjo no planejamento funcional: ao invés de continuar produzindo uma ilocução declarativa, a falante opta pela interrogativa, cuja forma exige algum tipo de reformulação. Essa é uma das discussões que sustenta a separação entre reparos de primeira e segunda ordens. A operação de *reformatting*, segundo Schegloff (2013), atua tanto para reparos *prima facie*, isto é, realizados do ponto de vista gramatical, quanto para aqueles que envolvem mudanças na função básica do trecho reparado.

O exemplo 15 mostra um caso que pode ser analisado como *replacing* - em operação de primeira ordem - quanto como *reformatting* - em operação de segunda ordem.

Exemplo 16 (Schegloff, 2013)

Vir: You want me to write you a a little list every week?

Mom: ***I would- (.) that would be great.***

Para Schegloff (2013), do ponto de vista puramente estrutural, trata-se de um caso de *replacing*, isto é, a substituição de *I* por *that*. Embora essa interpretação seja válida, classificá-lo apenas como *replacing* faria perder de vista o valor interacional da operação realizada pela falante. Nesse caso, o que ocorre é uma mudança de perspectiva: em vez de uma leitura centrada no desejo de Mom (*I would [love that]*), o foco se desloca para a sugestão de Vir (*that would be great*). A manutenção da perspectiva só poderia ser capturada na categorização desse fenômeno como *reformatting* (Schegloff, 2013).

A diferença entre reparos de primeira e segunda ordens é certamente um passo importante no reconhecimento de que as disfluências não se limitam a conjuntos de operações puramente estruturais. No entanto, argumentamos à frente neste trabalho que a ação de distinguir tipos de disfluências segundo suas diferenças de valor interacional não parece resistir a longos trabalhos de anotação. Do nosso ponto de vista, qualquer um desses tipos guarda motivações que interferem, em maior ou menor grau, na função interacional. Por esse motivo, pensamos ser mais acertada uma diferença entre tipos de operações e tipos de motivações. Esse assunto será abordado no capítulo de resultados.

2.1.2.10. Reordering

A operação de *reordering* ocorre quando o falante tenta reformular a ordem com a qual os elementos do turno foram previamente organizados. O exemplo 16 mostra um caso em que o falante reordena *get always* para *always get*.

Exemplo 17 (Schegloff, 2013)

Que: ***But do you get always- do you always get
called on?***

Segundo Schegloff, assim como ocorre com os outros casos de reparos, a operação de *reordering* pode ser interpretada tanto como reparo de primeira ordem quanto de segunda.

2.1.3. A proposta de Kosmala (2024)

A proposta de Kosmala (2024) é fruto de adaptações de classificações desenvolvidas em vários trabalhos anteriores. Para a descrição de sua classificação, é necessário introduzir o debate por trás da noção de disfluências enquanto obstáculos vs. ferramentas de planejamento durante a produção da fala. Ao mesmo tempo em que esses fenômenos são amplamente

frequentes em interações espontâneas, são relativamente raros os momentos em que eles suscitam problemas de interpretação ou que chamem a atenção dos interlocutores pelas suas ocorrências. Em última instância, podem também ser considerados elementos que marcam fluência e naturalidade em determinados contextos. Um exemplo prático desses casos são alguns geradores de voz por inteligência artificial, como o NotebookLM (Google, 2025) e o Google Duplex, nos quais pausas preenchidas, interjeições e diferentes tipos de reparo são acrescentados às interações entre os agentes conversacionais de modo a emular um diálogo em sua forma natural.

É nesse contexto que Kosmala (2024) recupera a noção de “fluencema” apresentada em Götz (2013). Em linhas gerais, o fluencema se configura como uma categoria abstrata baseada em várias propriedades da linguagem que condicionam a percepção de fluência durante o processo comunicativo. As propriedades que influenciam a percepção de fluência são muitas e podem ser divididas em três grandes grupos: fluencemas de produção (como *speech rate*, pausas preenchidas/silenciosas, repetições e marcadores discursivos), fluencemas de percepção (como entonação, diversidade lexical e sotaque) e fluencemas não verbais (como gestos e expressões faciais). Uma proposta de análise das disfluências que leve em consideração os três grupos pode ser apreciada em Götz (2013), embora ela não tenha caráter classificatório ou pretensão de servir como modelo para anotação dos fenômenos. Por outro lado, representa uma visão bastante completa a respeito da influência das disfluências em vários outros níveis linguísticos e extralinguísticos. O fluencema, no entanto, se encontra no centro da classificação (esta, sim, uma classificação) proposta por Kosmala (2024).

A proposta de anotação de Kosmala (2024) conta com nove categorias que marcam a percepção de fluência. Elas são divididas em fluencemas vocais (*vocal markers*), morfossintáticos (*morpho-syntactic markers*) e periféricos (*peripheral markers*).

- **VOCAL MARKERS (VOC):** são tipos de fluencema caracterizados por propriedades acústicas, como duração e características segmentais.
 1. **Pausas preenchidas (Filled pauses) (FP):** consistem na inserção de segmentos vocálicos longos e estáveis no fluxo da fala.
 2. **Prolongamentos (Prolongations) (PR):** consistem no alongamento de sílabas e palavras; se diferenciam de alongamentos com funções prosódicas específicas, como finalização de sintagmas entonacionais.

3. ***Pausas silenciosas (Unfilled pauses) (UP)***: consistem em períodos de silêncios durante a produção dos enunciados. As pausas consideradas por Kosmala (2024) são aquelas que possuem duração superior ou igual a 250 ms.
4. ***Autorreparos (Self-repairs) (SR)***: são correções realizadas pelo falante em sua própria fala quando há substituição de um trecho por outro. Podem ser distinguidos em (a) reparos sintáticos, (b) reparos morfológicos (c) reparos lexicais ou (d) adições (Levelt, 1983; Shriberg, 1994).
5. ***Auto-interrupções (Self-interruptions) (SI)***: consistem em enunciados interrompidos pelo próprio falante de modo que a sequência seguinte não seja composicional semântica e/ou sintaticamente com a anterior (interrompida). Em Kosmala (2024), essa noção engloba o conceito de *false-starts*.
6. ***Palavras truncadas (Truncated words) (TR)***: consistem na interrupção de uma palavra antes de sua articulação completa.
7. ***Repetições idênticas (Identical repetitions) (IR)***: consistem de repetições de itens ou de sequências linguísticas já pronunciadas. Uma distinção potencialmente importante é entre a repetição disfluenta e aquela com valor retórico. Em sua classificação, Kosmala (2024) não faz diferenciações funcionais desse tipo.

- ***PERIPHERAL MARKERS:***

8. ***Termos para edição explícita (Explicit editing terms) (EDT)***: consistem em palavras e expressões que tornam explícita algum tipo de edição realizada pelo falante na sequência linguística. Geralmente, assumem formas como “oops”, “sorry”, “I don’t know” ou “what’s the word”.
9. ***Sons não lexicais (Non-lexical sounds) (NL)***: consistem em vocalizações sem valor lexical que desempenham funções interacionais, como certas inspirações, bocejos, risadas e cliques.

As categorias propostas por Kosmala (2024) cobrem uma vasta gama de fenômenos de disfluências e são diferenciadas majoritariamente com base na composicionalidade semântica e morfossintática das formas envolvidas. Os fluencemas podem co-ocorrer na mesma sequência linguística: palavras truncadas (6), por exemplo, tendem a orbitar autorreparos (4), do mesmo modo que pausas preenchidas (1) tendem a co-ocorrer com pausas silenciosas (3).

A diferença proposta por Kosmala (2024) entre *Vocal* e *Peripheral markers* não parece ser de todo esclarecida, já que *non-lexical sounds* e *explicit editing terms* também são

caracterizados por propriedades acústicas. Embora não enunciada, a maior diferença entre elas parece ser seus papéis nos fenômenos de disfluências. Elementos como bocejos e risadas podem não ser interpretados pelos falantes como momentos de disfluência (e geralmente não são), mas sim como respostas paralinguísticas para determinadas situações ou mesmo com funções comunicativas. Em muitos casos, inclusive, esses itens possuem força ilocucionária e podem constituir um enunciado ou fazer parte de sequências acionais maiores. Os *explicit editing terms*, por outro lado, parecem reunir fenômenos de naturezas funcionais diversas, e a principal motivação de serem classificados dessa forma é justamente o fato de que eles ocorrem quando o falante se dá conta da disfluência cometida. Ao suspender momentaneamente o fluxo do enunciado e dizer “*what’s the word*”, ele manifesta ostensivamente que está em busca do melhor lexema para completar a sequência linguística. Esse fenômeno é análogo àquele que Schegloff (2013) chama de *parenthesizing*. Na perspectiva da *Language into Act Theory*, que será descrita na sequência, esses momentos não constituem disfluências *a priori*, mas alterações temporárias no aqui e agora da comunicação, realizadas num outro nível da fala: o nível parentético (Rocha, 2025). Esse recurso, no entanto, não se restringe à função de sinalização de reparos ou a momentos de disfluências.

2.1.4. A discussão terminológica

O assunto das disfluências está embebido em discussões terminológicas que são consequência de uma diversidade de nomeações utilizadas para os fenômenos estudados ao longo das últimas décadas. O mesmo pode ser observado para a identificação da área, que inclusive já foi nomeada pausologia (O’Connell & Kowal, 1980). Embora esta designação esteja intimamente relacionada com a ocorrência de pausas, seu uso é abrangente o suficiente para comportar fenômenos dos mais diversos.

A tabela 1 é adaptada de Kosmala (2024) e representa uma série de nomes bem estabelecidos para o estudo das disfluências. Cada um deles é respaldado por contextos teóricos de preocupações diversas. Os usos de cada um deles não são restritos aos autores selecionados. Não é incomum que o mesmo autor faça referência a essa área com uma variada gama de nomes, já que suas conceituações se sobrepõem na maior parte das vezes.

Tabela 1. Diferentes formas de se fazer referência às disfluências. Adaptado de Kosmala (2024).

Termo	Tentativa de definição	Autore(s)
-------	------------------------	-----------

Disfluência	Um desvio de uma forma ideal de realização da fala. Envolve uma suspensão temporária ou uma interrupção do fluxo.	Shriberg (1994), Eklund (2004), Lickley (2015), Ferreira & Bailey (2004)
Hesitação	Extensão temporal da mensagem (através de pausa ou algum tipo de atraso), geralmente associado à incerteza do falante ou a uma ação de escolha.	Gilquin (2008), Maclay & Osgood (1959), Duez (1991), Betz (2020)
Confluência	Co-construção da fluência durante a conversação	McCarthy (2009)
(Dis)fluência	Fenômeno ambivalente que pode funcionar como marca de fluência ou disfluência	Götz (2013), Crible et al. (2019)
Manutenção comunicativa	Fenômeno linguístico ou não que fornece evidência de que o falante está fazendo a manutenção da comunicação levando em consideração o reconhecimento do interlocutor.	Allwood et al. (1990), Ginzburg & Poesio (2016)
Wellformedness	Dispositivos utilizados pelo falante para produzir sequências linguísticas bem formadas.	Hieke (1981)
Travail de formulation	Marcas utilizadas para planejamento e formulação.	Morel & Danon-Boileau (1998)
Collateral signals	Sinais comunicativos que tem como função comentar e fazer a manutenção da performance do falante.	Clark (1996)
Inter-(dis)fluência	Fluencemas com funções ambivalentes de fluência e disfluência a depender da estrutura interacional e cognitiva.	Kosmala (2024)

O mesmo (ou talvez mais grave) problema terminológico ocorre na designação dos tipos de disfluências. Em um trabalho realizado com vinte e três anotadores experientes e com *backgrounds* variados, Trouvain et al. (2025) mostraram o papel da variabilidade das anotações realizadas para as mesmas sequências linguísticas. Os anotadores divergiam em diferentes dimensões. Para alguns casos, categorias de disfluência foram utilizadas para fenômenos de diferentes naturezas, como marcadores discursivos e ruídos paralinguísticos (deglutição, inspiração e etc.). Do mesmo modo, termos diferentes foram empregados com variados níveis de granularidade para os mesmos fenômenos, como *non-lexical filler*, *filler particle* e/ou *filled pause*. As principais divergências ocorreram em casos de reparos e de cadeias com disfluências sequenciais. Enquanto alguns anotadores desmembraram as partes do reparo (*reparandum*, *interregnum* e reparo) em diferentes *tiers*, outros simplesmente marcaram toda a cadeia de disfluências em sequência com a etiqueta de reparo.

O problema terminológico está longe de ser restrito ao estudo das disfluências. Por enquanto, a melhor maneira de controlá-lo é definir de forma evidente cada uma das categorias a serem utilizadas no trabalho e seus métodos de anotação. Em alguma medida, a instabilidade de conceitos e as discordâncias relevantes apenas apontam para a natureza multifacetada dos fenômenos e um longo caminho a ser percorrido pela área.

2.2. Language into Act Theory


A *Language into Act Theory* (L-AcT) é uma teoria *corpus-driven* construída principalmente pelo Laboratório de Italianística da Universidade de Florença (LABLITA), então coordenado por Emanuela Cresti e Massimo Moneglia. Na atual conjuntura, uma série de trabalhos lançam mão da L-AcT para a descrição e a compreensão da fala espontânea. Muitos desses trabalhos advêm do Laboratório de Estudos Empíricos e Experimentais da Linguagem (LEEL), da UFMG, através do uso de vários corpora de fala espontânea, como o C-ORAL-BRASIL I (Raso & Mello, 2012), o C-ORAL-ESQ (Rocha et al., 2025) e seus respectivos minicorpora.

A L-AcT é apresentada como uma das consequências teóricas da Teoria dos Atos de Fala (Austin, 1962) com extensa fundamentação empírica para o estudo das ilocuções e da estrutura informacional. A teoria conta com suas próprias definições acerca das unidades básicas da fala em face de propriedades prosódicas e de seu caráter acional. Tanto a ilocução quanto a fronteira prosódica terminal ocupam papéis centrais nessas definições.

São duas as unidades de referência que governam a organização informacional e possibilitam a interpretabilidade pragmática na fala: os enunciados e as stanzas. Eles se diferenciam principalmente com relação à quantidade de atos ilocucionários veiculados em uma sequência linguística que se encerra com uma quebra prosódica terminal (a sequência terminada). Ambos são unidades de referência que contam com autonomias prosódica - são sequências terminadas - e pragmática. Os enunciados dizem respeito a sequências terminadas prosodicamente que veiculam uma única ilocução. As stanzas, por sua vez, são consideradas sequências terminadas que veiculam duas ou mais ilocuições.

As sequências terminadas - tanto em enunciados quanto em stanzas - podem ser realizadas por diferentes unidades entonacionais que estabelecem entre si diferentes tipos de sinais de continuidade prosódica até a fronteira terminal. Essas unidades prosódicas estão relacionadas sistematicamente a unidades informacionais que enriquecem textual e funcionalmente o ato de fala. A ilocução é veiculada por uma ou por um conjunto de sílabas nas chamadas unidades informacionais ilocucionárias. Ao redor delas, o falante tem a possibilidade de realizar outras unidades com diferentes funções informacionais. O conjunto da unidade ilocucionária com as outras unidades funcionais opcionais é chamado de padrão informacional. A fala, desse ponto de vista, passa a ser compreendida como uma troca de atos de fala, e a estrutura informacional atua enriquecendo o conteúdo locutivo e garantindo a interpretabilidade da unidade ilocucionária, única efetivamente necessária para o ato de fala.

O exemplo 18 representa um enunciado realizado através de uma única unidade entonacional. Nesse caso, essa unidade é responsável por veicular o conteúdo semântico e a função ilocucionária. As barras duplas “//” representam uma fronteira prosódica terminal, marcando o fim da sequência terminada. As barras simples “/”, como no exemplo 19, marcam fronteiras não terminais.

Exemplo 18 bfamnn01_3341 C-ORAL-BRASIL I (dB-COM) 

**MAI: enrolou toda no corpo dele //*

A onda sonora e a curva de frequência fundamental desse exemplo se encontram na figura 1. Em uma única unidade prosódica terminada (autonomia prosódica), o falante veicula o conteúdo locutivo e imprime a força ilocucionária (autonomia pragmática). Esse é o exemplo mais simples de padrão informacional, composto por uma única unidade ilocucionária.

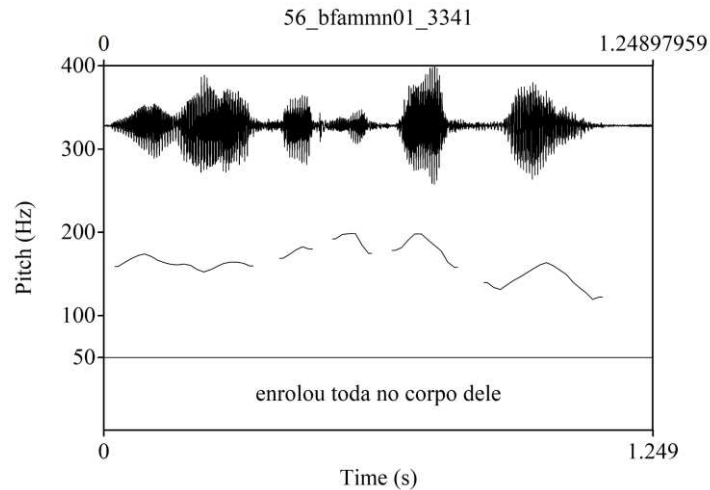


Figura 1. Frequência fundamental e forma de onda do exemplo 17.

O exemplo 18 se contrasta com os exemplos 19 e 20, que representam, respectivamente, um enunciado composto por três unidades informacionais - sendo uma ilocucionária (ILL) e duas não ilocucionárias (IU) - e uma stanza, composta por duas unidades ilocucionárias e seis unidades não ilocucionárias (IU).

Exemplo 19 *bpubd101_3855 C-ORAL-BRASIL I (dB-COM)* 🎧

*PAU: pedreiro /^{IU} nũ tem muita dúvida não /^{ILL} né //^{IU}

As unidades entonacionais de 19 podem ser visualizadas na figura 2. Nesse exemplo, é possível perceber que as unidades possuem contornos prosódicos distintos, da mesma forma que veiculam diferentes funções informacionais. A descrição acerca de contornos entonacionais já havia sido bem documentada no âmbito do IPO (*Institute for Perception Research/Eindhoven*) (t' Hart; Cohen; Collier, 1990), e a relação das unidades prosódicas à estrutura informacional também contava com direções apontadas em Chafe (1970), Halliday (1976) e outros. No entanto, a associação sistemática de diferentes formas prosódicas a diversas funções informacionais é uma das inovações da L-AcT.

Em 19, a primeira unidade, “peDREIro”, tem a função de estabelecer uma representação semântica - nesse caso, evocar um grupo de entidades genéricas - com relação à qual a locução dirá respeito (*pragmatic aboutness*). A esta unidade informacional, a L-AcT dá o nome de Tópico¹¹. Sua forma entonacional, visualizada na figura 2, parece ter se convencionalizado na

¹¹ O panorama geral da unidade de Tópico e de todas as outras unidades informacionais identificadas pela L-AcT será apresentado nos subtópicos seguintes.

estrutura da linguagem para veicular essa função. O Tópico está relacionado à capacidade de *displacement* (Hockett, 1960); um enunciado sem ele necessariamente se refere ao contexto.

O mesmo raciocínio pode ser aplicado para as outras unidades informacionais não ilocucionárias¹², como é o caso daquela composta pela interjeição “né”. Diferentemente do Tópico, essa unidade não acrescenta conteúdo semântico ao enunciado, mas possui a função de auxílio dialógico, i.e., se trata de um marcador discursivo. O padrão informacional nesse caso é dado pelo conjunto da unidade ilocucionária e as duas unidades não ilocucionárias adjacentes.

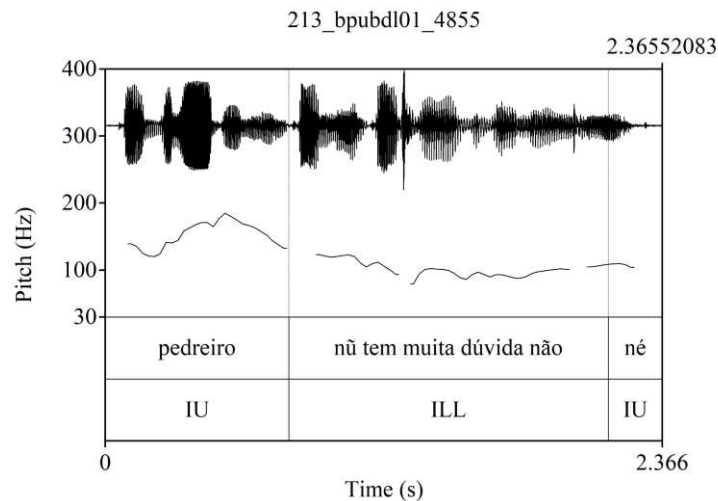


Figura 2. Frequência fundamental e forma de onda do exemplo 18.

O exemplo 20 representa uma sequência terminada relativamente longa em comparação com as duas anteriores. Além de produzir seis unidades informacionais não ilocucionárias, o falante realiza duas ilocuções, na penúltima e na última unidades. Trata-se de uma stanza. Entre uma unidade ilocucionária e outra, não há fronteira prosódica terminal. Ao invés disso, o falante finaliza a primeira unidade ilocucionária com um sinal de continuidade prosódica. Apenas após a segunda ilocução é que se percebe a terminalidade. São, portanto, duas ações realizadas subsequentemente na mesma sequência terminada. Do mesmo modo, são dois padrões informacionais realizados na mesma sequência terminada. Esses casos ocorrem quando o falante está mais interessado em elaborar os conteúdos semântico e informacional que contextualizam as ilocuções.

¹² O caso é um pouco diferente para as unidades ilocucionárias, já que a ilocução depende de parâmetros pragmático-cognitivos para além do núcleo prosódico realizado em uma unidade entonacional.

Exemplo 20 *bfamnn04_3624 C-ORAL-BRASIL I (dB-COM)* 🎧

*REG: aí eu /^{IU} assim /^{IU} depois que passou tudo /^{IU} né /^{IU} cê fica sozim /^{IU} em plena enfermaria /^{IU} Haroldo ainda ficou um tempo lá comigo /^{ILL} e foi levar o Gabriel //^{ILL}

A identificação de fronteiras prosódicas terminais e não terminais é realizada através de uma série de pistas acústicas para além do comportamento de f_0 (pausas, mudanças bruscas de intensidade, *reset* de f_0 , taxa de elocução, etc.). A segmentação é feita por falantes competentes em treinamento constante. O acordo entre os segmentadores é aferido através do teste *Kappa* (Mello et al., 2012).

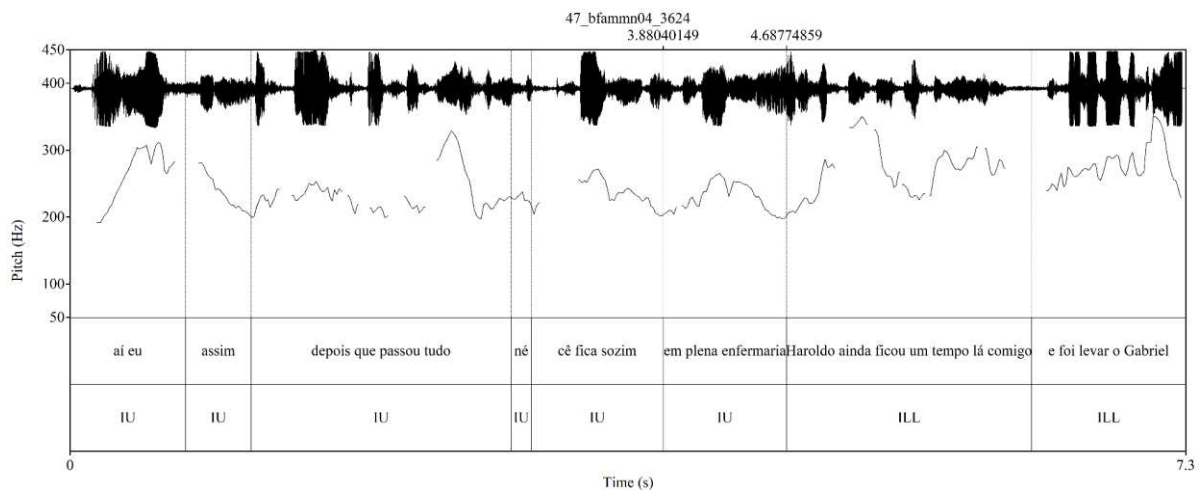


Figura 3. Frequência fundamental e forma de onda do exemplo 19.

São quatorze unidades informacionais identificadas pela L-AcT. Três delas são de tipo textual ilocucionário, cinco são textuais não ilocucionárias e seis são unidades dialógicas. A anotação depende de três parâmetros básicos: função da unidade, forma prosódica e posição com relação à unidade ilocucionária. Cada unidade ilocucionária pode estabelecer seu padrão informacional, composto por unidades não ilocucionárias que orbitam a ilocução.

Assim como a proposta original de Austin (1962), a L-AcT propõe uma divisão tripartida de atos linguísticos que decorrem da execução do ato de fala. O **ato locucionário** (ou locutivo) diz respeito ao conteúdo semântico e à estrutura morfossintática dos enunciados e das stanzas. O **ato ilocucionário** (ou ilocutivo) se relaciona às ações intencionalmente dirigidas pelo falante ao seu interlocutor (asserção, ordem, pergunta, etc.). Enquanto o ato locucionário é codificado linguisticamente pela semântica e pela morfossintaxe, a ilocução possui sua contraparte estrutural veiculada principalmente pela prosódia das unidades ilocucionárias. O **ato perlocucionário** é aquele em que a L-AcT mais se diferencia da formulação austiniana. Ele

é formulado como uma força pulsional, resultado da relação afetiva entre falante e interlocutor. Nesse sentido, é a perlocução que faz com que o falante possua a intenção de realizar o ato de fala. Essa interpretação é fundamentada na chamada *Human Birth Theory* (Fagioli, 1972) e pode ser melhor compreendida em Cresti (2025).

No próximos subtópicos, serão apresentadas as unidades informacionais utilizadas contemporaneamente para anotação de corpora de fala no âmbito da L-AcT.

2.2.1. As unidades informacionais

As unidades informacionais se dividem em dois grandes grupos: as unidades textuais e as unidades dialógicas. Ao grupo das unidades textuais pertencem aquelas que acrescentam conteúdo proposicional ao enunciado e às stanzas. Elas se subdividem em unidades textuais ilocucionárias - que veiculam ilocuições - e não ilocucionárias. Ao segundo grupo, das unidades dialógicas, pertencem aquelas que não acrescentam conteúdo proposicional e nem são composicionais semântica ou morfossintaticamente com o restante da sequência linguística. As dialógicas são dirigidas aos interlocutores e têm a função de regular a interação.

2.2.1.1. As unidades textuais

As unidades textuais são aquelas que acrescentam conteúdo semântico à sequência linguística. Elas são divididas em textuais ilocucionárias e textuais não ilocucionárias. As unidades textuais ilocucionárias são o Comentário (COM), o Comentário Ligado (COB) e o Comentário Múltiplo (CMM). As unidades textuais não ilocucionárias são o Tópico (TOP), o Apêndice de Tópico (APT), o Apêndice de Comentário (APC), o Introdutor Locutivo (INT) e o Parentético (PAR).

2.2.1.1.1. Comentário (COM)

O Comentário é a unidade informacional responsável por veicular a força ilocucionária. Ele conta com uma proeminência prosódica (núcleo) cuja forma é um elemento importante para determinar a ilocução a ser realizada. A força ilocucionária não é dependente de toda a forma prosódica da unidade, mas apenas de seu núcleo, realizado por meio de um pequeno número de sílabas. As propriedades prosódicas do núcleo variam em função da ilocução. O Comentário é a única unidade necessária para compor um enunciado. Ao redor dele, podem ser encontradas as outras unidades que formam o padrão informacional.

A classificação do tipo de força ilocucionária não depende apenas da forma prosódica do núcleo do Comentário, mas também de parâmetros pragmático-cognitivos. Isso ocorre pelo

fato de que a quantidade de formas prosódicas é menor do que o número de ilocuções possíveis de serem realizadas. A identificação das ilocuções depende de metodologias experimentais que permitem separar aspectos pragmático-cognitivos, valores atitudinais e a própria forma do núcleo da unidade (Rocha, 2013; Raso & Rocha, 2017). A L-AcT conta com sua própria classificação ilocucionária, que se diferencia daquelas apresentadas preliminarmente em Austin (1962) e em Searle (1975).

2.2.1.1.2. Tópico (TOP)

A unidade de Tópico (TOP) tem a função de fornecer o domínio para a interpretação da força ilocucionária. O TOP fornece representações semânticas - de natureza espacial, temporal, individual etc. - que devem ser levadas em consideração para a interpretação da ilocução. Além dos Comentários, TOP é a única unidade que conta com núcleo prosódico que se convencionalizou para expressar essa função informacional. Ele se encontra sempre à esquerda da unidade ilocucionária para a qual foi realizado.

As unidades de TOP podem ser realizadas a partir de três formas prosódicas diferentes (Cavalcante, 2020):

- Tipo 1
Seu núcleo está posicionado na última sílaba tônica e nas pós-tônicas (caso haja). A f₀ possui movimento ascendente-descendente e pode haver alongamento de sílaba.
- Tipo 2
Seu núcleo está posicionado na última sílaba tônica e nas pós-tônicas (caso haja). A f₀ possui movimento ascendente e pode haver alongamento de sílaba.
- Tipo 3
Há dois semi-núcleos: o primeiro possui valor alto ou extra-alto de f₀, enquanto que o segundo possui valor mais baixos. Pode haver a presença de alongamento silábico.

2.2.1.1.3. Apêndice de Comentário (APC)

A unidade informacional e Apêndice de Comentário (APC) possui a função de acrescentar conteúdo semântico e material morfossintático ao comentário. As informações veiculadas pelo APC geralmente correspondem a conteúdos dados ou facilmente recuperáveis pelo contexto. O APC não possui núcleo prosódico, e sua forma entonacional pode ser achatada

ou descendente. Ele se localiza sempre à direita da unidade de comentário. Em virtude de sua função e de suas propriedades prosódicas, a omissão do APC não compromete a interpretabilidade pragmática da sequência linguística.

2.2.1.1.4. Apêndice de Tópico (APT)

O Apêndice de Tópico (APT) possui função análoga à do APC, isto é, acrescentar conteúdo semântico e material morfossintático à unidade de tópico. Ele está posicionado à direita do TOP, e sua forma pode ser nivelada, ascendente ou, em alguns casos, reproduzir de maneira abrandada a forma do tópico, sem a proeminência funcional.

2.2.1.1.5. Parentético (PAR)

A unidade de Parentético (PAR) possui função de acrescentar explicações ou fornecer ao interlocutor instruções pragmáticas acerca de como interpretar o conteúdo semântico realizado no padrão informacional. Os PAR geralmente são utilizados com funções metalinguísticas, metanarrativas ou modalizadoras. São realizados com média de f0 abaixo do nível do enunciado, sem núcleos prosódicos. Geralmente possuem taxa de articulação e intensidade diferentes daquelas desempenhadas nas outras unidades informacionais.

O trabalho de Rocha (2025) sistematiza observações importantes a respeito dos chamados parentéticos curtos e longos. Os primeiros têm sido compreendidos como uma unidade informacional de fato, cujas propriedades foram descritas no parágrafo anterior. Neles, ocorrem tipicamente uma quantidade limitada de expressões linguísticas (“assim”, “por exemplo”, “tipo”, “eu acho”, “por aí” e “aqui”) (Santos & Bossaglia, 2018). Os parentéticos longos, por outro lado, podem ser interpretados como um outro nível discursivo, realizado quando o falante suspende temporariamente o aqui e o agora do enunciado e realiza padrões informacionais que são subordinados ao nível discursivo ilocucionário (para uma discussão detalhada desse assunto, ver Rocha, 2025).

O padrão distribucional dos parentéticos é livre, com a exceção do início absoluto do enunciado. Na anotação, a unidade informacional Parentético, anotada como PAR, se diferencia do nível parentético, no qual a anotação é realizada considerando as unidades informacionais realizadas no nível subordinado (PAR-TAG, em que TAG pode dizer respeito a qualquer unidade informacional). Em contraste com o *parenthesizing*, descrito por Schegloff (2013), o nível parentético na L-AcT possui função pragmática mais abrangente e não está ligado necessariamente à sinalização ostensiva de uma hesitação.

2.2.1.1.6. Introdutor Locutivo (INT)

A unidade de Introdutor Locutivo (INT) tem a função de sinalizar que haverá algum tipo de suspensão do aqui e do agora do enunciado, introduzindo uma metailocução, como é o caso do discurso reportado. Quanto à forma, o INT é descrito como possuindo f0 descendente e alta taxa de articulação em comparação com o restante da sequência. O INT sempre precede o padrão informacional que é por ele introduzido.

2.2.1.1.7. Comentário Múltiplo (CMM)

Em determinadas sequências terminadas, o falante realiza ilocuções que possuem relação prosódica convencionalizada entre si. Esse é o caso do Comentário Múltiplo (CMM), unidade ilocucionária que, seguida de outro(s) CMM(s), estabelece um padrão melódico convencional que possibilita uma interpretação holística das ilocuções veiculadas. Nas cadeias de CMM, é possível perceber que as ilocuções seguintes - que fazem parte do padrão melódico - foram planejadas antes mesmo da primeira unidade que compõe esse padrão. Em outras palavras, o planejamento das ilocuções em Comentários Múltiplos não se dá de maneira processual, como é o caso dos COBs (apresentados na próxima subseção), mas de modo padronizado.

Diferentes padrões ilocucionários estabelecem relações funcionais diversas, como CMMs alternativos, chamamento funcional, comparação, lista, relação necessária e reforço. Não raro, as unidades de CMM também podem estabelecer relações hierárquicas entre si, quando uma ilocução é subordinada a outras ilocuções do mesmo padrão (Mendes, 2024).

2.2.1.1.8. Comentário Ligado (COB) e a stanza

O Comentário Ligado (COB) ocorre em sequências em que o falante realiza mais de uma ilocução sem que elas estabeleçam entre si uma relação padronizada melodicamente ou convencionalizada funcionalmente. É, portanto, uma unidade ilocucionária e, diferentemente do COM e do CMM, conta com sinal de continuidade, i.e., não possui autonomia prosódica. As ilocuções veiculadas em sequências de COBs tendem a ser enfraquecidas e homogêneas, além de pertencerem a uma mesma classe ilocucionária (Moneglia & Raso, 2014).

A stanza é uma das unidades de referência formada por ao menos um Comentário Ligado e um Comentário. As unidades de COB podem definir o seu próprio padrão informacional, assim como o COM. Elas são facilmente encontradas em monólogos, em que o falante possui maior controle do turno conversacional e maior capacidade de planejamento de longas sequências terminadas.

As stanzas são importantes para a comparação da complexidade informacional em corpora cujas disfasias são incomparáveis *a priori*. A metodologia apresentada em Rocha et al. (2022) propõe uma comparação baseada em stanzas de mesmo número de unidades ilocucionárias. Dessa forma, sequências com duas ilocuções (COB e COM) em um corpus são comparadas a sequências com duas ilocuções (também COB e COM) em outro corpus. O mesmo é feito para sequências com outras quantidades de ilocuções. Essa técnica permite analisar estruturas que possuem o mesmo potencial de complexidade informacional, minimizando vieses gerados pela incomparabilidade contextual. Para trabalhos que utilizam essa metodologia, veja Raso et al. (2023), para o C-ORAL-ESQ, e Rocha & Vital (2025), para o C-ORAL-BRASIL II (Raso et al., 2025).

2.2.1.2. As unidades dialógicas

As unidades dialógicas, também conhecidas como marcadores discursivos, são dirigidas ao interlocutor e têm a função de regular a interação de variadas maneiras. Elas não são composicionais semântica ou sintaticamente com o restante da sequência linguística e são encontradas em unidades entonacionais dedicadas. Além disso, não podem ser interpretadas em isolamento, além de não possuírem autonomia prosódica. Um panorama geral das unidades dialógicas pode ser encontrado em Raso (2014) e em Santos (2024), em diferentes etapas de sua compreensão no contexto da L-AcT.

O Incipitário (INP) possui a função de iniciar um padrão informacional. Ele pode ser de dois tipos: o primeiro dá início ao subpadrão enquanto estabelece um contraste afetivo (não dependente do conteúdo lexical) com o que foi dito anteriormente. Nesse caso, sua forma prosódica conta com alta variação de f_0 e curta duração. O segundo tipo de INP dá início ao padrão sem estabelecer o contraste com a sequência anterior. Sua forma prosódica é nivelada (*flat* INP).

O Conativo (CNT) tem a função de sinalizar a resolução ilocucionária do enunciado. Sua forma prosódica é marcada por um movimento descendente de f_0 alinhado à vogal tônica. Em alguns casos, observa-se um movimento ascendente de f_0 em vogais anteriores à vogal tônica (preparação).

O Alocutivo (ALL) geralmente ocorre em posição final do padrão informacional - com algumas ocorrências em posição medial. Sua forma prosódica é marcada por um movimento descendente seguido de um nivelamento de f_0 e baixa intensidade em comparação com outras unidades. É comum que ocorram apagamentos de segmentos na posição de nivelamento de f_0 .

O Expressivo (EXP) possui a função de estabelecer coesão social com o interlocutor, fornecendo o suporte emocional para a unidade ilocucionária. Santos (2024) propõe uma definição mais específica, afirmando que o EXP é utilizado para expressar surpresa durante a realização do padrão informacional e é realizado de modo que não possa ser confundido com uma unidade ilocucionária. Sua forma prosódica é marcada por um movimento ascendente de f_0 até a vogal tônica. A sílaba tônica também é caracterizada por um alongamento, em comparação com a duração das sílabas do comentário. Geralmente, o EXP ocorre no início do padrão informacional.

O Evidenciador (EVD) ou *Highlighter* (HGL) possui a função de dirigir a atenção do interlocutor para o conteúdo realizado anteriormente. Sua forma prosódica é caracterizada por um movimento de subida rápida de f_0 , baixa intensidade e curta duração em comparação com o COM.

O Conector Discursivo (DCT) possui a função de marcar a continuidade entre enunciados ou entre padrões informacionais. São frequentemente realizados por itens dessemantizados, como conjunções e preposições. Sua forma prosódica é variável, motivo pelo qual se discute sua relação com as escansões, que serão tratadas em breve neste texto..

2.2.2. Sintaxe e macrossintaxe na L-AcT

Diferentemente do padrão informacional, a sintaxe não ocupa um papel central na organização da fala na L-AcT. O enunciado e a stanza - enquanto sequência que podem ser formadas por diferentes unidades informacionais - não são construídos primariamente a partir de relações sintáticas, mas funcionais. O papel da sintaxe é descrito de maneira subordinada às relações informacionais marcadas prosodicamente.

Com o estudo de corpora orais, é possível notar a maior ocorrência de estruturas sintáticas mais simples em comparação com aquelas presentes em corpora de escrita. Em sua proposta original, Cresti (2000) sugere que o âmbito da sintaxe da fala é limitado às unidades informacionais, que funcionam como ilhas; as relações entre essas unidades, por outro lado, dependem principalmente das relações funcionais que elas estabelecem entre si através da prosódia. Em casos de enunciados simples, isto é, formados por uma única unidade informacional, podem ou não haver relações sintáticas entre os itens linguísticos que formam o conteúdo locutivo. Em enunciados complexos, formados por mais de uma unidade informacional multpalavras, a composicionalidade sintática é interpretada dentro dos limites da unidade. Dessa forma, não é necessário que as relações sintáticas extravasem as fronteiras informacionais - embora isso possa ocorrer em alguns casos, como na subordinação e

coordenação -, pois a composicionalidade entre elas é de natureza pragmática e realizada por outro mecanismo linguístico: a prosódia.

Uma das consequências dessa proposta é que um padrão informacional pode existir mesmo sem que a sintaxe opere entre os itens do conteúdo locutivo. Isso ocorre pois as relações informacionais são primárias na construção da sequência linguística. A articulação entre as unidades informacionais e as relações sintáticas, caso existam, ocorre num nível posterior de construção na fala: a chamada macrossintaxe (Cresti, 2000; Blanche-Benveniste, 2003).

Para uma visão panorâmica do papel da sintaxe na L-AcT, ver Lima e Silva (2020) e Bossaglia (2015). Da nossa parte, é importante perceber o papel primário da construção informacional sobre as relações sintáticas. São, na verdade, as relações informacionais que definirão as relações sintáticas da fala, e não o contrário. Esse ponto precisa ser considerado na descrição das disfluências apresentada neste trabalho. Como propomos na sequência, as operações de disfluência atuam sobre variados níveis linguísticos (segmental, morfossintático, semântico e pragmático). Em casos de reparo morfossintáticos, por exemplo, são focalizadas as estruturas no interior da unidade informacional. Por outro lado, casos de reparos informacionais ou ilocucionários naturalmente extrapolam os limites informacionais.

2.2.3. As disfluências na L-AcT

O tratamento dado às disfluências na *Language into Act Theory* varia a depender do momento de desenvolvimento da teoria e dos corpora levados em consideração. No âmbito do *Corpus di Italiano Parlato* (Cresti, 2000), utilizado para a primeira apresentação sistemática da L-AcT, os fenômenos de disfluência são descritos partindo da discussão acerca de sua representação no sistema de transcrição para a fala espontânea. O sistema adotado nesse corpus é o CHAT (McWhinney, 2000). A partir de algumas adaptações, foram definidos os critérios de transcrição para os corpora C-ORAL-ROM (Cresti & Moneglia, 2005) e C-ORAL-BRASIL I (Raso & Mello, 2012). Nosso objetivo nesta seção é descrever as particularidades e os fenômenos de disfluência anotados nos três sistemas de transcrição.

2.2.3.1. Anotações no *Corpus di Italiano Parlato*

No *Corpus di Italiano Parlato* (Cresti, 2000), as disfluências são tratadas sob o conceito de fenômenos de fragmentação da fala, compreendidos não como elementos causados por acidentes externos, mas instâncias naturais do processo de verbalização. Os fenômenos são descritos em dois grandes grupos: aqueles que ocorrem no ataque do enunciado (os *retractings* ou falsas-partidas) e aqueles que ocorrem na sua conclusão (as interrupções). Já nas primeiras

descrições, se deixam claras as intenções de diferenciar os tipos de fragmentação de suas causas em potencial (*esitazione, cambio di programma, interruzione di un programma*, etc.). A diferenciação entre os dois grupos em Cresti (2000) - *retractings* e interrupções - ocorre em função de suas consequências diversas na constituição do enunciado.

Os *retractings* são compreendidos como fenômenos causados pelo processo hesitativo que pode ocorrer quando o falante verbaliza os primeiros elementos da sequência linguística. Em outras palavras, ele indica uma indecisão (ou incerteza) do falante a respeito de como dar início à verbalização. Essa indecisão, segundo Cresti (2000), precede a realização do enunciado, motivo pelo qual o *retracting* é classificado como disfluência de ataque.

Os *retractings* são classificados em três tipos que se diferenciam com relação à duplicação do conteúdo locutivo envolvido na disfluência: *retractings* de duplicação total, de duplicação parcial e sem duplicação.

O *retracting* de duplicação total ocorre quando todo o conteúdo locutivo de uma unidade entonacional é duplicado na unidade seguinte, após o fenômeno de fragmentação. O conteúdo realizado na primeira unidade - e repetido na segunda - não é dotado de funcionalidade informacional. Esse caso é sinalizado no *Corpus di Italiano Parlato* como “[/]", conforme o exemplo 21.

Exemplo 21 Corpus di Italiano Parlato (Cresti, 2000)

*GNO: eh ! **c' è [/] c' è / la [/] la** strisciata //

O *retracting* de duplicação parcial ocorre quando apenas uma parte do conteúdo locutivo da unidade entonacional é duplicada na próxima unidade. A anotação ocorre com [//], conforme o exemplo 22.

Exemplo 22 Corpus di Italiano Parlato (Cresti, 2000)

*IMP: però / **c' è la [//] c' è** scritto sulla circolare / che / al momento in cui vengono erogati i soldi / deve comunque essere presentato <il contratto> //

É fundamental notar que a definição de *retracting* de duplicação parcial em Cresti (2000) não exige que o trecho não duplicado seja composicional com a unidade subsequente. No caso 22, a preposição “la” não foi duplicada, pois passou simplesmente por um processo de deleção. A principal preocupação é distinguir o fenômeno em face de repetição total, parcial ou em termos de não haver repetição do conteúdo locutivo.

O *retracting* sem duplicação ocorre quando nenhum item da unidade entonacional envolvida na fragmentação é repetido na unidade seguinte. Sua anotação se dá com o símbolo “[///]”, como no exemplo 23.

Exemplo 23 *Corpus di Italiano Parlato* (Cresti, 2000)

*PNI: aimmeno &vo [///] spero / d' andare un po' in giro //

Os fenômenos de *retracting* devem ser distintos daquelas fragmentações que ocorrem na conclusão dos enunciados (Cresti, 2000), a saber, as interrupções. Nelas, o falante não está hesitante com relação à escolha do programa informativo; ao contrário, esse programa já foi escolhido e está em curso. O que ocorre é que o falante toma a decisão de não o terminar. Eventualmente, pode haver um processo de replanejamento da sequência interrompida, mas este não é compreendido como fruto de um procedimento hesitativo para a escolha de alternativas para o ataque do enunciado, como no caso dos *retractings*. As interrupções são marcadas com “+”, como no exemplo 24.

Exemplo 24 *Corpus di Italiano Parlato* (Cresti, 2000)

*PN3: son tornata + perché / prima di quell' ora / in casa / 'un mi riceveva / vero ?

Um terceiro fenômeno descrito por Cresti (2000) é a interrupção do programa informativo, anotada “+?”. Esta constitui a maior parte das interrupções e ocorre quando há uma mudança da estratégia informativa. O conteúdo locutivo realizado até a interrupção é reinterpretado por parte do interlocutor com base no princípio de cooperação (Grice, 1975). Exemplo 25 representa uma interrupção do programa informativo.

Exemplo 25 *Corpus di Italiano Parlato* (Cresti, 2000)

*SAB: sicchè / i tifosi incavolati +/ a questo punto c' hanno fatto scedere / siamo andati per le strade di Milano / fino allo stadio / a piedi / sotto l' acqua //

Outros fenômenos também são sinalizados, mas suas relações com as disfluências não são tão claras. A chamada suspensão intencional ou interrupção intencional (“...”) ocorre quando o falante interrompe o curso de seu próprio enunciado que já foi planejado para que seja suspenso. Esse caso se assemelha ao chamado perfil suspensivo, numa interpretação contemporânea da L-AcT. A chamada hesitação sem interrupção (“-?”) ocorre quando o falante realiza mudança de unidade entonacional sem mudar a unidade informacional. Esse também

parece ser um fenômeno que tem sido reinterpretado contemporaneamente como uma simples escansão.

A tabela 2 representa os tipos de fenômenos discutidos no âmbito do *Corpus di Italiano Parlato* como exemplos de disfluências.

Tabela 2. Tipos de fenômenos anotados como disfluência no *Corpus di Italiano Parlato*.

Fenômeno	Anotação
Retracting de duplicação	[/]
Retracting de duplicação parcial	[//]
Retracting sem duplicação	[///]
Hesitação sem interrupção	-?
Interrupção causada pelo interlocutor	+/.
Interrupção intencional (suspensão)	...
Interrupção para escolha locutiva ou informativa	+
Interrupção do programa informativo	+/

Segundo Cresti (2000), não há fenômeno de retracting em que não ocorra mudança de unidade entonacional, do mesmo modo em que não há interrupção em que não ocorra mudança de unidade informacional. Esse detalhe já sugere uma preocupação em propor uma anotação de disfluências pautada na continuidade ou não da estrutura informacional. A validade empírica dessa proposição, no entanto, ainda não havia sido sistematicamente verificada.

2.2.3.2. Anotações no C-ORAL-ROM

O C-ORAL-ROM (Cresti & Moneglia, 2005) é um corpus de fala espontânea que documenta as principais línguas românicas faladas na Europa: italiano, francês, espanhol e português europeu. Boa parte dos critérios de transcrição e de segmentação prosódica são inspirados no sistema CHAT e nas potenciais alterações realizadas para o *Corpus di Italiano Parlato*.

A documentação dos fenômenos de disfluência ou dos fenômenos de fragmentação no âmbito no C-ORAL-ROM é resumida em comparação à discussão realizada em Cresti (2000). São mantidos os três tipos de *retractings* (com uma nova interpretação para o *retracting* sem

duplicação); com exceção da suspensão intencional do enunciado (que não é mais considerada fenômeno de fragmentação), todos os tipos de interrupções passam a ser identificados na mesma anotação (“+”). Se mantém a noção de que sempre que houver um *retracting* haverá uma quebra prosódica não terminal, ao passo que as interrupções geram sempre quebras prosódicas com valor terminal.

As pausas preenchidas são anotadas como “&he”, independentemente do segmento produzido. No âmbito informacional, elas são anotadas como TMT (*time taking units*).

A tabela 3 resume as propostas do C-ORAL-ROM.

Tabela 3. Tipos de fenômenos anotados como disfluências no C-ORAL-ROM.


Fenômeno	Anotação
Interrupção do enunciado causada por mudança na programação pelo falante ou interrupção externa	+
Retracting com repetição total	[/]
Retracting com repetição parcial	[//]
Ambiguidade retracting/interrompido	[///]
Pausa preenchida	&he (independentemente do segmento)

A anotação “[///]” que marcava o *retracting* sem repetição de palavras em Cresti (2000) passa a ser compreendida em Cresti & Moneglia (2005) como uma ambiguidade entre *retracting* e interrupção. Ainda assim, a escolha entre os diferentes tipos de *retractings* se dá com base na reutilização ou não de itens linguísticos ditos anteriormente. Esse critério se diferencia significativamente da interpretação do *retracting* no C-ORAL-BRASIL.

2.2.3.3. Anotações no C-ORAL-BRASIL

O C-ORAL-BRASIL I (Raso & Mello, 2012) é o corpus de referência do português brasileiro falado informalmente. Sua arquitetura, seus processos de transcrição e de segmentação foram planejados de modo a garantir a comparabilidade com o C-ORAL-ROM. Por conta disso, os critérios de transcrição foram inspirados naqueles utilizados pelo corpus europeu, com alterações importantes para capturarem fenômenos em gramaticalização e lexicalização no português brasileiro (Mello et al., 2012). As disfluências anotadas também consistem em *retractings* e em enunciados interrompidos.

Diferentemente do C-ORAL-ROM, os *retractings* no C-ORAL-BRASIL seguem uma anotação baseada no número de palavras que foram suprimidas pelo falante após o fenômeno de disfluência. O símbolo utilizado é “[/n]”, em que “n” representa o número de palavras da unidade entonacional que não fazem parte do texto final da sequência linguística. O exemplo 26 mostra um caso em que apenas uma palavra é alvo de retratação por parte do falante.

Exemplo 26 *bfamcv04 C-ORAL-BRASIL I* 

*BRU: cê não pode aprontar [/1] apontar pra mesa //

Essa representação - com o número de palavras que não fará parte do texto final do enunciado - possui suas vantagens nas etapas seguintes da compilação do corpus, como a subtração do total de palavras e a anotação morfossintática. Nesta última, as palavras imediatamente à esquerda do *retracting* (cuja quantidade é indicada pelo numeral) são desconsideradas para a anotação PoS ou para as anotações de funções sintáticas e de dependências. Embora possuam funções importantes, este esquema de anotação acaba por reduzir os fenômenos de disfluência envolvidos nessas unidades ao simples cancelamento de palavras realizadas pelo falante equivocadamente.

A distinção realizada no C-ORAL-ROM entre *retractings* de repetição total e de repetição parcial acaba sendo recuperada no C-ORAL-BRASIL através da anotação informacional. Caso todas as palavras da unidade tenham sido repetidas na unidade após o *retracting*, utiliza-se a etiqueta EMP (*empty*). Caso algum/alguns dos itens da unidade em que houve a retratação seja composicional com a unidade seguinte, utiliza-se a etiqueta SCA (*scanning unit*).

Assim como no C-ORAL-ROM, as pausas preenchidas são identificadas na transcrição por “&he”, independentemente das propriedades acústicas do segmento realizado. Na etiquetagem informacional, as unidades tonais realizadas por pausas preenchidas são anotadas com a sigla TMT.

Tabela 4. Tipos de fenômenos anotados como disfluências no C-ORAL-BRASIL I.

Fenômeno	Anotação
Retracting	[/n]
Sequência interrompida	+
Escansão	SCA (no nível da etiquetagem informacional)

Pausa preenchida	&he (independentemente do segmento)
------------------	-------------------------------------

3. Metodologia

Este capítulo apresenta a metodologia deste trabalho. A investigação se dividiu em três partes: (1) o desenvolvimento da classificação, (2) a anotação do corpus e (3) a análise dos dados. Todas elas foram realizadas com base em corpora da família C-ORAL, principalmente o C-ORAL-BRASIL I (Raso & Mello, 2012), o C-ORAL-BGEST (Barros & Mello, 2025) e o C-ORAL-ROM (Cresti & Moneglia, 2005). A classificação criada foi utilizada para anotação de uma parte do minicorpus do C-ORAL-BRASIL (Bossaglia; Cavalcante; Raso, 2018), a partir do qual os dados também serão analisados. Esse minicorpus já conta com alinhamento texto-som e anotação informacional segundo a *Language into Act Theory*.

Este capítulo está organizado da seguinte forma: inicialmente será apresentado o corpus C-ORAL-BRASIL I e seu minicorpus etiquetado; na sequência, serão descritos os procedimentos realizados para a criação da classificação de disfluências. A última parte será dedicada aos métodos empregados na análise quantitativa das disfluências em face das anotações de estrutura informacional já presentes nos textos do corpus.

3.1. C-ORAL-BRASIL I

O principal corpus utilizado neste trabalho é o C-ORAL-BRASIL I (Raso & Mello, 2012), corpus de referência do português brasileiro falado informal. Compilado entre 2008 e 2012, o C-ORAL-BRASIL I conta com 139 gravações de fala espontânea transcritas semi-ortograficamente e segmentadas em unidades entonacionais. Todas as gravações contam com alinhamento texto-som realizado através do software WinPitch (Martin, 2010). Para a totalidade dos textos, foi realizada etiquetagem morfosintática através do parser PALAVRAS (Bick, 2012). A transcrição e a segmentação contam com materiais de validação estatística (Mello et al., 2012), garantindo acurácia com relação aos critérios definidos e acordo estatístico entre os transcritores/segmentadores. O corpus está integralmente disponível para download no site do grupo C-ORAL-BRASIL¹³ e também pode ser acessado pela interface DB-CoM¹⁴ para consultas na rede.

A arquitetura do C-ORAL-BRASIL é baseada principalmente na alta variação diafásica. Esse parâmetro é um dos mais importantes para o registro de uma gama variada de ações linguísticas condicionadas por situações interativas diversificadas. Os textos do corpus contam com uma subdivisão pautada no contexto interativo e na tipologia interacional. Os contextos

¹³ <https://www.c-oral-brasil.org/corpora_para_download.php>

¹⁴ <<http://150.164.100.23/>>

são o público ou o privado, e eles se diferenciam principalmente com relação ao papel dos participantes na interação, isto é, se eles interagem como representantes de uma instituição (possuem papel profissional) ou como indivíduos. O contexto privado representa 75% das interações do corpus, contra 25% de interações em contextos públicos. Cada contexto possui gravações em diferentes tipologias interacionais (monólogos, diálogos e conversações) que se diferenciam com relação à quantidade de participantes em interação efetiva.

O número de falantes é 362, dos quais 247 (68,23%) foram coletadas informações sociodemográficas, como sexo, idade, origem e escolaridade. O número total de palavras é 208.130, com uma média de 1.500 palavras por texto. O contexto familiar/privado conta com 159.364 palavras, enquanto que o contexto público possui 48.766 palavras. O impacto dos falantes cujas informações não foram coletadas é pequeno, já que eles pronunciam até 47 palavras cada (apenas 1,9% da totalidade das palavras do corpus) (Raso, 2012).

O balanceamento entre os grupos sociodemográficos é feito através da quantidade de palavras pronunciadas pelos falantes de cada uma das categorias (sexo, idade e escolaridade). O balanceamento da categoria de sexo é praticamente perfeito, com 50% das palavras do corpus sendo pronunciadas por indivíduos do sexo masculino e 50% por indivíduos do sexo feminino. A idade foi dividida em 6 classes apresentadas a seguir com seus respectivos percentuais de palavras pronunciadas:

- M (menor de 18 anos): 1,6% de palavras
- A (entre 18 e 25 anos): 27,1% de palavras
- B (entre 26 e 40 anos): 30,3% de palavras
- C (entre 41 e 60 anos): 31,0% de palavras
- D (maior de 60 anos): 8,1% de palavras
- X (idade desconhecida): 1,9% de palavras

A categoria de escolaridade é considerada com base no tempo de escolarização e no exercício da profissão. Foram definidos 4 grupos, representados a seguir com a percentagem de palavras pronunciadas:

- Faixa 1 (ausência de escolarização ou escolarização até o nível primário incompleto): 15,78% de palavras
- Faixa 2 (escolarização até o terceiro grau completo, com exercício de profissão que não exija esse título): 40,76% de palavras
- Faixa 3 (terceiro grau completo e exercício de profissão que exija esse título): 40,66% de palavras
- Faixa X (nível escolar desconhecido): 2,79% de palavras

Com essa documentação, o corpus representa a variedade padrão do português na diatopia mineira. Embora não represente a maior parte das palavras pronunciadas, é expressiva também a representação da variedade popular (Faixa 1 de escolaridade).

Outras informações de interesse sociolinguístico podem ser apreciadas no cabeçalho de cada gravação. Se trata de um arquivo .txt que acompanha cada um dos textos, fornecendo as informações sociodemográficas de cada participante e detalhes técnicos (duração, tópico, situação, data, qualidade acústica, etc.) e linguísticos (número de palavras, formas aferéticas, formas convencionalizadas nos critérios de transcrição, erros de pronúncia, etc.). O cruzamento automático desses dados do cabeçalho com aqueles disponíveis nas transcrições e nos arquivos de anotação morfosintática vem sendo cada vez mais viabilizado pelas técnicas de Processamento de Língua Natural e da Sociolinguística Computacional (Rocha, 2024).

3.1.1. O minicorpus etiquetado informacionalmente

O C-ORAL-BRASIL conta com um minicorpus (Bossaglia; Cavalcante; Raso, 2018) de 20 textos etiquetados informacionalmente segundo a *Language into Act Theory* (L-Act). Para a constituição do minicorpus, foram selecionados 15 textos em contexto familiar/privado e 5 em contexto público nas diferentes tipologias (monólogo, diálogo e conversação), conforme a tabela 5.

Tabela 5. Número de textos no minicorpus segundo as tipologias e os contextos.

		Contexto	
		Familiar	Privado
Tipologia	Monólogo	6	1
	Diálogo	5	2
	Conversação	4	2

Uma série de critérios foram utilizados na seleção dos textos do C-ORAL-BRASIL I para comporem o minicorpus. A principal preocupação foi criar uma amostra de gravações que pudesse ser representativa da totalidade do corpus, de modo que os resultados obtidos em análises do minicorpus pudessem ser expandidos para toda a parte informal do C-ORAL-BRASIL (os 139 textos). Para além da subdivisão de tipologias e contextos, os textos

selecionados contam com alta variabilidade de atividades e situações comunicativas, qualidade acústica suficiente para análise entonacional durante todo o trecho das gravações e balanceamento de informações sociodemográficas.

Todos os textos do minicorpus foram etiquetados em unidades informacionais segundo a L-AcT. Diferentemente da anotação morfossintática, que pode ser realizada de maneira automatizada através dos parsers disponíveis, a etiquetagem informacional é feita manualmente por etiquetadores humanos em treinamento contínuo para essa finalidade. O minicorpus está disponível em duas versões que se diferenciam com relação às atualizações no âmbito das unidades informacionais descritas pela L-AcT: a versão de 2012 e a de 2018. Nosso trabalho lançou mão apenas da versão mais atual, aquela de 2018.

Para o trabalho com as disfluências, foram selecionados dez textos do minicorpus, cobrindo cada uma das tipologias e dos contextos. Optamos por essa amostragem - ao invés do uso do minicorpus inteiro - em virtude do alto volume de disfluências, cujas anotações exigem tempo e revisão atenta para evitar inconsistências. Os textos selecionados estão dispostos na tabela 6, com informações técnicas relativas à gravação e à situação comunicativa. No total, são documentados 15 falantes do sexo masculino e 17 do sexo feminino. O número de palavras é 15.923, próximo da metade do total de palavras do minicorpus completo. Os textos variam em duração, totalizando 01:28:24 (hh:mm:ss).

Tabela 6. Gravações selecionadas do minicorpus para a anotação de disfluências.

Texto	Título	M	F	Palavras	Duração
bfamcv01	Campeonato de futebol	4	0	1482	00:07:00
bfamcv02	Casamento em família	0	3	1736	00:07:51
bfamdl01	Compras no mercado	0	4	2150	00:14:39
bfamdl02	Guardando equipamentos	1	1	1581	00:07:26
bfammn01	A cobra	2	0	1085	00:05:02
bfammn02	Drummond	1	1	1692	00:07:23
bfammn03	Dívida	3	3	1211	00:07:08
bpubcv01	Doação de sangue	2	3	1805	00:08:30
bpubdl01	Construção	2	0	1563	00:14:09
bpubmn01	Língua estrangeira na escola pública	0	2	1616	00:09:16


Total	15	17	15.923	01:28:24
-------	----	----	--------	----------

O número de gravações selecionadas em relação aos contextos e às tipologias é comparável ao número disponível no minicorpus: sete gravações provenientes de contexto familiar privado e três provenientes de contexto público. A tipologia melhor representada são os monólogos, com quatro gravações. Como já apresentado no capítulo 2, os monólogos condicionam estruturas textuais mais complexas, uma vez que o falante tem a chance de elaborar melhor as sequências linguísticas sem muitas preocupações em perder o turno conversacional. Os dados são balanceados com relação à tipologia (mn: 5.604 palavras; dl: 5.294 palavras; cv: 5.023 palavras). Há, no entanto, um desequilíbrio com relação ao contexto (familiar: 10.937 palavras; público: 4.984 palavras). Os resultados que dependem do tamanho dos textos serão normalizados para uma quantidade de 1000 palavras. Para análise da estrutura informacional, também será utilizada normalização para cada 100 unidades informacionais.

Originalmente, o alinhamento texto-som do minicorpus foi realizado através do software WinPitch (Martin, 2004). As gravações foram recentemente convertidas de forma manual para o software Elan (Wittenburg et al., 2006), *software* que tem recebido atualização e suporte constantes. A versão de 2018 do minicorpus está disponível para download nas duas versões de alinhamento (pelo WinPitch e pelo Elan) no site do C-ORAL-BRASIL. As anotações das disfluências foram realizadas todas pelo Elan.

3.2. A anotação das disfluências

A anotação das disfluências foi feita diretamente nos arquivos de alinhamento do Elan de cada uma das gravações. Às trilhas de transcrição, foram acrescentadas etiquetas entre dois sinais de arroba (“@”) marcando o tipo de disfluência e sua motivação, separados por um traço *underline* (“_”). A transcrição, as anotações de retractings e as etiquetas informacionais foram todas mantidas. O exemplo 25 ilustra a anotação de um reparo (*repair*) motivado pela tentativa de melhorar o conteúdo semântico do enunciado (*improvement*). A anotação nesse caso é realizada com a *tag* “@repair_improvement@” localizada logo após a interrupção do enunciado por parte do falante. O trecho após a *tag* já representa o enunciado com as devidas reelaborações.

Exemplo 27 bpubd101 Minicorpus C-ORAL-BRASIL I 

*ROG: eu vou &coloc [/3]=EMP=@repair_improvement@ eu vou suspender mais um pouquim aqui /=CMM= vou pegar a linha /=CMM= e vou colocar por cima //CMM=

Optamos pela não criação de uma trilha separada para a anotação das disfluências, uma vez que nem todos os fenômenos analisados são essencialmente dependentes do tempo, isto é, são melhor representados de modo pontual, sem duração definida. Na mesma trilha, há as etiquetas informacionais, que serão utilizadas para a descrição da posição das disfluências em face das partes constitutivas do padrão informacional. Dessa forma, poderemos saber de forma automática em qual unidade ocorreu o fenômeno de disfluência e quantificar seus tipos e motivações em face da estrutura informacional.

A maior preocupação na etapa metodológica, no entanto, foi definir um conjunto fechado de tipos e motivações de disfluências que pudesse ser depreendido dos dados por indução - daí o trabalho de natureza *corpus-driven*. A definição dessas classes se deu em conjunto com os orientadores em dois momentos permeados por diferentes níveis de reflexão teórica: em meados de 2024 e em meados de 2025. A proposta de 2024, documentada em Vital; Rocha; Raso (2025), propõe a existência de seis classes de disfluências que determinariam a continuidade ou o abandono da unidade informacional que estava sendo construída quando o fenômeno ocorria: reparo, repetição, mudança de planejamento, escansão, pausa preenchida e abandono. As motivações catalogadas à época foram as seguintes: prospecção, incerteza, problema de articulação, resposta imediata, melhoramento textual, satisfação e interrupção externa. Nessa classificação, os tipos de disfluência foram considerados determinantes para que o falante optasse pelo abandono ou pela continuidade da unidade informacional em desenvolvimento.

A proposta de 2025, que será melhor elaborada aqui, utiliza praticamente os mesmos rótulos (com alguns acréscimos e algumas reelaborações) para tipos e motivações. Entretanto, percebemos que as motivações - e não os tipos - é que parecem determinar a continuidade ou a fragmentação das unidades informacionais. Com essa interpretação, deixam de existir tipos específicos de disfluência que condicionam alterações informacionais e tipos que não as condicionam; na realidade, a atual classificação permite pensar que quaisquer tipos de disfluência podem gerar alterações tanto no nível da estrutura informacional quanto no nível da estrutura locutiva. O que determina essas alterações são precisamente as motivações dos fenômenos e não seus tipos. A elaboração formal desta proposta será abordada detalhadamente

no próximo capítulo. Neste momento, é importante conceber a ideia de que nossa classificação se vale de disfluências que podem ocorrer em dois níveis simultaneamente: o nível locutivo, relativo à estrutura linguística, morfossintática e semântica do enunciado, e o nível informacional-ilocucionário, de caráter pragmático, relativo ao padrão informacional, à ilocução e à atitude. Uma disfluência ocorrida no nível informacional-ilocucionário pode desencadear outras disfluências no nível locutivo, enquanto que o inverso parece não ocorrer. Esses pontos serão esclarecidos na seção de resultados. Como etapas metodológicas, a definição e a conceituação das classes serão apresentadas a seguir.

3.2.1. As classes

A classificação desenvolvida por meio do estudo *corpus-driven* comporta sete tipos de disfluências e oito motivações. Tanto o grupo dos tipos quanto o grupo das motivações contam com uma etiqueta para casos incertos (*unclear*). As classes são as seguintes.

TIPO DE DISFLUÊNCIA

- ***Repair (reparo)***: o reparo ocorre quando o falante realiza uma autocorreção no conteúdo ou na forma gramatical daquilo que foi dito. No nível locutivo, são atestados reparos segmentais, morfossintáticos, semânticos e lexicais; no nível informacional-pragmático, são atestados reparos de ilocução, de estrutura informacional e de atitude.
- ***Repetition (repetição)***: a repetição ocorre quando o falante repete o(s) mesmo(s) item(ns) sem associar valor funcional ao ato de repeti-los. Em alguns casos, o item repetido é realizado com articulação ou formas diferentes, mas o referente se mantém. Em casos de repetições em sequência, os itens são anotados individualmente como *repetition*. As repetições parecem ser mais frequentes no nível locutivo e podem ser consequência de disfluências no nível informacional que impactam a estrutura do enunciado.
- ***Change of planning (mudança de planejamento)***: a mudança de planejamento ocorre quando o falante faz uma mudança na maneira de apresentar a informação. A informação substancial (ou o conteúdo proposicional) se mantém (quasi-)inalterada, e o que muda é apenas seu modo de apresentação ao interlocutor. Na anotação, diferenciamos dois tipos de mudança de planejamento:
 - ***Change of informational planning (mudança de planejamento informacional)***: quando há mudança no planejamento do nível informacional (alterando a composição do padrão informacional). Esses casos podem gerar

outras disfluências do conteúdo locutivo, já que o falante precisará determinar outra locução que veicule o pensamento da melhor forma possível.

- ***Change of textual planning (mudança de planejamento locutivo)***: quando há mudança no planejamento da estrutura linguística empregada para veicular a informação. Nesse caso, o falante faz alterações nas estruturas morfossintáticas e/ou semânticas para a acomodação do conteúdo informativo. Mudanças no nível locutivo parecem não impactar o programa informacional.
- ***Scanning (escansão)***: as escansões ocorrem quando o falante realiza uma unidade informacional em duas ou mais unidades entonacionais (quebra do isomorfismo). Em outras palavras, o falante produz uma escansão quando ele coloca uma quebra prosódica não terminal antes que se encerre a unidade informacional em curso. Casos em que isso ocorre com propósitos retóricos não são considerados como disfluências. A quebra do isomorfismo através das escansões pode gerar outros fenômenos no nível locutivo, como as repetições de palavras ou pausas preenchidas.
- ***Filled pause (pausa preenchida)***: as pausas preenchidas ocorrem quando o falante suspende o fluxo de sua fala ou está prestes a iniciá-lo e preenchem parte do turno com segmentos geralmente vocálicos e/ou nasais. As pausas preenchidas não alteram a escolha informacional ou a estrutura locutiva, mas fornecem tempo para que ambas possam ser planejadas da melhor forma possível.
- ***Abandonment (abandono)***: o abandono ocorre quando o falante abandona por completo a sequência linguística que estava sendo construída. Essa ação ocorre volitivamente por variados motivos.

MOTIVAÇÕES

- ***Prospection (prospecção)***: a prospecção ocorre quando o falante produz uma disfluência com o objetivo de prospectar/procurar pela melhor forma de se expressar ou pelo melhor programa informacional.
- ***Uncertainty (incerteza)***: a incerteza está relacionada com a hesitação na escolha de um programa informacional para ser seguido ou de uma estrutura locutiva para ser utilizada. Geralmente ocorre quando o falante não está certo do conteúdo semântico que está prestes a ser veiculado.
- ***Articulation problem (problema de articulação)***: o problema de articulação representa um erro na pronúncia de qualquer estrutura locutiva. Geralmente a articulação imprecisa

leva o falante a realizar outras disfluências, como reparos, repetições e pausas preenchidas. O problema de articulação é típico do nível locutivo.

- ***Inadequacy (inadequação)***: a inadequação se dá quando o falante realiza uma estrutura locutiva ou um programa informacional e, na sequência, a(o) considera inadequada(o) para expressar o conteúdo desejado. A inadequação não ocorre apenas do ponto de vista gramatical, mas também do lexical, como em escolhas de palavras imprecisas ou incorretas em face daquilo que o falante deseja comunicar.
- ***Improvement (melhoramento)***: o melhoramento ocorre quando o falante realiza algum tipo de disfluência com o objetivo de melhorar ou de tornar o discurso mais claro. Ele pode ocorrer no nível informacional - quando há a disfluência para melhorar o programa informacional - ou no nível locutivo - quando a disfluência ocorre para melhorar o texto do enunciado.
- ***Program alteration (alteração do programa)***: quando há alteração do programa informacional ou textual que estava sendo executado pelo falante.
- ***External interrupt (interrupção externa)***: a interrupção externa ocorre quando o falante se interrompe para que sua fala não se sobreponha a um evento externo qualquer. Nesse caso, não há intenção de resposta imediata por parte do falante.
- ***Immediate response (resposta imediata)***: a resposta imediata ocorre quando o falante se interrompe com o objetivo de produzir uma resposta imediata a um evento externo, como uma pergunta ou um comentário do interlocutor). A resposta do falante ao interlocutor pode não ser de natureza linguística, mas de natureza puramente acional.
- ***Satisfaction (contentamento)***: o contentamento ocorre quando o falante produz uma disfluência (geralmente a mudança de planejamento ou o abandono) pelo fato de já ter alcançado seus objetivos comunicativos antes do fim da sequência linguística. Por já se encontrar satisfeito (contentado) com a realização da sequência parcial, ele produz um abandono ou muda seu planejamento.

Tanto o tipo de disfluência quanto a motivação contam com uma etiqueta de casos em que não é possível identificar uma dentre as diferentes classes. Esses casos são marcados como *unclear* e geralmente ocorrem quando não há material linguístico suficiente para decidirmos com precisão entre uma classe e outra.

3.3. A quantificação

Após a anotação de todos os dez textos selecionados do minicorpus, foram extraídos do ELAN os arquivos .txt com codificação UTF-8 em formato “texto de transcrição tradicional”. Esses arquivos contam com a identificação dos falantes, a transcrição, tempo de início e fim das anotações e os silêncios maiores ou iguais a 100ms. O exemplo 26 a seguir representa o trecho do arquivo bpubcv01, que ilustra a estrutura dos dados extraídos do software de alinhamento.

Exemplo 28 bpubcv01 C-ORAL-BRASIL I 

```

001 Flávia      o' /=CNT= Bruno /=ALL= a gente vai explicar como é que
a gente trabalha aqui //COM=
                00:00:00.058 - 00:00:02.596
002 Emmanu     <isso> //COM=
                00:00:02.596 - 00:00:03.166
003 Bruno      <ham ham> //COM=
                00:00:02.596 - 00:00:03.166
004 Bruno      <hum hum> //COM=
                00:00:03.166 - 00:00:04.013
005 Flávia     essa <daqui /=TOP= é a> bolsa de sangue /=COB=
igualzinho ela chega da coleta //COM=
                00:00:03.166 - 00:00:06.375
006 Bruno      ham ham //COM=
                00:00:06.375 - 00:00:06.888
                (0.46)
007 Flávia     sai do doador /=CMM= lá /=APC= e chega pra gente assim
//CMM=
                00:00:07.344 - 00:00:09.134

```

Todos os arquivos .txt extraídos do ELAN com as anotações informacionais e as anotações de disfluência estão disponíveis no domínio <<https://osf.io/c84mx/files/osfstorage>>. A quantificação dos fenômenos se deu através de um script em Python que estrutura os dados de cada um dos .txt's em um *dataframe*. Com os dados estruturados, foram contabilizadas as anotações de tipos e motivações de disfluências, além de suas posições com relação às unidades

informativos. Tanto o dataframe quanto o script podem ser acessados no domínio especificado acima.

4. Resultados

4.1. A classificação *corpus-driven*

Conforme apresentado na seção de metodologia, a etapa de desenvolvimento da classificação resultou em 7 categorias de disfluência e 8 categorias de motivação. Cada uma delas será apresentada e exemplificada nesta seção. Optamos por manter os nomes em inglês e português. Em muitos casos, alternamos no texto entre um e outro. Todos os exemplos deste capítulo contam com arquivo de áudio disponibilizado nos paratextos da dissertação.


4.1.1. As categorias de disfluência

Foram sete as categorias de disfluência utilizadas na anotação: *repair* (reparo), *scanning* (escansão), *repetition* (repetição), *abandonment* (abandono), *filled pause* (pausa preenchida), *change of textual planning* (mudança de planejamento textual) e *change of informational planning* (mudança de planejamento informacional). Elas foram definidas através de análises empíricas, mas também possuem algumas similaridades com as categorias previamente apresentadas por Schegloff (2013) e Kosmala (2024). As categorias podem atuar sobre a estruturação do enunciado (conteúdo locutivo) ou sobre suas funções pragmáticas (conteúdo informacional-ilocucionário). Nesta apresentação, as propriedades de cada uma são descritas em função da distinção entre conteúdo locutivo e informacional-ilocucionário. Esta distinção insere a proposta no âmbito da *Language into Act Theory*.

Todos os exemplos apresentados a partir de agora contam com arquivo de áudio numerado. O download do repositório completo pode ser feito no repositório da dissertação no Open Science Framework. Este link também se encontra no Anexo A desta dissertação: <https://osf.io/c84mx/files/osfstorage>

4.1.1.1. *Repair* (reparo)


O reparo ocorre quando o falante realiza uma autocorreção, isto é, repara o conteúdo ou a forma de alguma expressão linguística que foi completa ou parcialmente realizada por ele. Essa operação pode atuar sobre qualquer nível da língua: reparos fonéticos corrigem pronúncias mal executadas, enquanto reparos morfossintáticos atuam sobre as variadas estruturas morfológicas e sintáticas. Um reparo semântico pode ocorrer quando o significado evocado por uma forma linguística é diferente daquele que se deseja comunicar. O reparo pragmático envolve os níveis mais altos do significado, como a ilocução, a estrutura informacional e a atitude.

Exemplo 29 bfamcv02 C-ORAL-BRASIL I 

*RUT: e a [/2]=EMP=@repair_inadequacy@ e o meninim /=COM= né //EVD=

No exemplo 29, a falante realiza na primeira unidade entonacional apenas a conjunção e o artigo feminino: “e a”. O reparo ocorre precisamente pelo fato de que o gênero selecionado para o artigo deveria ser masculino, de modo a concordar com o núcleo “meninim”. Em autocorreções como essa, é bastante frequente que o falante repita determinadas estruturas realizadas antes do reparo, mesmo sem que elas estejam envolvidas na reformulação. Esse é o caso da conjunção “e”, que foi repetida na unidade entonacional seguinte. Podemos classificar esse reparo como de natureza morfossintática, realizado em virtude de a falante perceber como inadequada a não concordância de gênero que estava prestes a ser realizada: “a mininim”.


O exemplo 30, da mesma gravação, representa um caso de reparo lexical, no qual a falante seleciona incorretamente o lexema “padrinho”. Nesse caso, a percepção de inadequação ocorreu já durante a realização da palavra, que é prontamente interrompida após a primeira sílaba “pa”.

Exemplo 30 bfamcv02 C-ORAL-BRASIL I 

*RUT: pelo amor de Deus /=COB= nũ me convida pa ser &pa
[/1]=SCA=@repair_inadequacy@ madrinha não /=COB= hein //COM=

Os dois exemplos acima dizem respeito a reparos que ocorrem no nível do conteúdo locutivo dos enunciados. Neles, o falante comete imprecisões de natureza estrutural ou lexêmica que não impactam a função pragmática das unidades em que eles ocorrem. No exemplo 29, se o reparo não fosse realizado, e a falante pronunciasse “e a meninim” ao invés de “e o meninim”, essa unidade prosódica manteriam sua função informacional (se trata de um Comentário) e a mesma força ilocucionária seria realizada. O mesmo parece ser o caso no reparo lexical do exemplo 30, em que dizer “padrinho” ao invés de “madrinha” não impactaria a função informacional nem alteraria a força ilocucionária da unidade. Em outras palavras, a mesma unidade informacional que estava sendo realizada antes do reparo é aquela que ocorre após a autocorreção. Essa análise, no entanto, não é válida para o exemplo 31, a seguir. Nele, a disfluência não atua alterando o conteúdo locutivo, mas a própria ilocução. O trecho

pronunciado de maneira fragmentada “&aq” parece em vias de veicular uma força declarativa¹⁵. A unidade seguinte, no entanto, veicula uma pergunta.

Exemplo 31 *bfamd102* *C-ORAL-BRASIL I* 

*BRU: não / &aq [/1]=EMP=@repair_program alteration@ aqui // =COM=

O reparo exemplificado em 31 não deve ser encarado como de mesmo nível dos anteriores, uma vez que a alteração não é realizada sobre o conteúdo locutivo da unidade, mas sobre sua função ilocucionária. Um caso semelhante pode ser observado no exemplo 32, em que o reparo ocorre no nível informacional. Nesse exemplo, o falante produz uma lista cujos elementos são os nomes de três times de futsal. Cada elemento é produzido como uma unidade ilocucionária. O reparo em questão ocorre entre a primeira e a segunda unidades “Squadra / é Squadra”. A primeira unidade se trata de um elemento de uma lista processual - realizada como Comentário Ligado (COB) - enquanto que as unidades seguintes são padronizadas - realizadas como Comentários Múltiplos (CMM). A alteração realizada é de natureza informacional, pois o falante apenas substituiu um COB por um CMM, duas unidades ilocucionárias.

Exemplo 32 *bfamcv01* *C-ORAL-BRASIL I* 

*GIL: Squadra [/1]=EMP=@repair_program alteration@ é Squadra / =CMM= José Mourinho / =CMM= Amarelão // =CMM=

Outro reparo que pode ser interpretado como de nível pragmático é o do exemplo 33. Observando apenas o texto do enunciado, é possível pensar que não se trata de um reparo, mas de uma simples repetição de palavras “como é / como é”. A escuta do áudio, no entanto, se torna essencial para percebermos que o conteúdo locutivo repetido foi realizado de duas maneiras prosodicamente diferentes, conforme a figura 4.

Exemplo 33 *bfammn02* *C-ORAL-BRASIL I* 

*LUC: como é [/2]=EMP=@repair_inadequacy@ como é que começou // =COM=

Nossa proposta interpretativa para esse caso é a de que houve um reparo atitudinal. O falante realiza a primeira unidade (“como é”) com uma atitude menos marcada, que pode ser visualizada no movimento de f0 da primeira unidade entonacional. O trecho em que ele repete

¹⁵ Não é nosso objetivo dizer se a ilocução foi ou não realizada nesse caso. O mais importante é notar que houve uma mudança de ilocução: de um ato potencialmente declarativo para uma pergunta.

essa estrutura possui movimento prosódico similar, embora com média de f_0 maior do que o anterior.

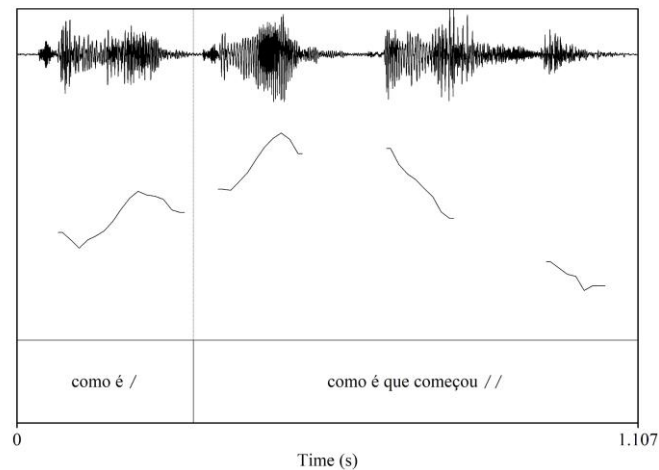


Figura 4. Forma de onda e f_0 do exemplo 31.


4.1.1.2. *Scanning* (escansão)

As disfluências chamadas *scannings* ou escansões ocorrem quando o falante realiza uma unidade informacional em duas ou mais unidades entonacionais. A anotação das escansões na L-AcT já segue o paradigma apresentado no capítulo 2. A anotação de *scanning* para as disfluências, por outro lado, engloba uma quantidade maior de fenômenos - todos relacionados com a inserção de uma quebra prosódica não terminal antes da finalização de uma unidade informacional.

O exemplo 34 conta com uma unidade de TOP (“se a gente vai de táxi”) que acabou sendo produzida em duas unidades informacionais: “se a gente vai de” e “de táxi”. A inserção da quebra prosódica não terminal antes da finalização do tópico também faz com que a falante repita a preposição “de”, que finaliza a unidade escansionada e também inicia a unidade que realiza prosodicamente a forma de TOP. Essa repetição não constitui a disfluência primária desse trecho, mas se trata simplesmente de uma consequência da *scanning*. Geralmente, as palavras repetidas nesses casos são itens funcionais, e a repetição pode ser uma tentativa de manter suas relações sintáticas nos limites de uma mesma unidade entonacional (conforme 2.2.2.).

Exemplo 34 *bfamd101* C-ORAL-BRASIL I 

*FLA: o' /=COB= se a gente vai de [/1]=SCA=@scanning_prospection@ de táxi /=TOP=
voltar de táxi /=PAR= po' comprar um //COM=

Exemplo 35 *bfamcv02* C-ORAL-BRASIL I 

*RUT: da família &d [/3]=EMP=@scanning_unclear@ <da família> dele /=TOP= minha
filha /=ALL= <também nũ> tomo nem conhecimento /=COM= também //APC=

Exemplo 36 *bfamcv01* C-ORAL-BRASIL I 

*GIA: tipo /=INT= eu [/1]=EMP=@scanning_unclear@ eu acredito /=i-COM= tipo /=PAR= cem
por cento nisso //COM=

Os exemplos 35 e 36 contam com *scannings* em unidades textuais. Ambos são marcados tradicionalmente como *retractings* que envolvem o cancelamento da totalidade de palavras da unidade entonacional (*retractings* do tipo EMP). As mesmas palavras canceladas (“da família &d”, no exemplo 33, e “eu”, no exemplo 34) são repetidas na unidade entonacional seguinte. À primeira vista, esses casos poderiam ser compreendidos na nossa classificação como repetições, mas não o são. Parece razoável a possibilidade de que as repetições só ocorreram pelo fato de que os falantes, por motivos desconhecidos, inseriram a quebra não terminal antes do fim da unidade informacional. É precisamente essa inserção que caracteriza o fenômeno de *scanning*: trata-se de um erro de execução (Cresti, 2000) bastante frequente de natureza prosódica. Parece ser o caso então que o desmembramento da unidade informacional em diferentes unidades entonacionais leva a outras disfluências aparentes, como as repetições.

4.1.1.3. *Repetition* (repetição)

As repetições ocorrem quando o falante repete, sem função comunicativa, trechos ditos anteriormente. Diferentemente da *scanning*, para que ocorram repetições, não é necessária a quebra do isomorfismo entre uma unidade informacional e uma unidade entonacional. Da mesma forma, a repetição não exige quebra prosódica entre os trechos repetidos. No exemplo 37, há repetição da contração “dum”, realizada no interior de uma unidade ilocucionária. Entre o primeiro e o segundo item não há quebra prosódica e nem variação formal que justifique a repetição como um elemento dotado de valor funcional.

Exemplo 37 bfamnn01 C-ORAL-BRASIL I 

*MAI: nũ sei /=COB= né /=AUX= a imaginação hhh dum dum @repetition_unclear@
animal /=COM= o que que pode ser /=APC= né //AUX_n=

Exemplo 38 bfamnn01 C-ORAL-BRASIL I 

*MAI: só ni lugar /=SCA=@scanning_prospection@ que tem tem @repetition_articulation problem@
mata mata @repetition_unclear@ muito /=SCA=@scanning_prospection@ forte /=COB= mata
/=SCA=@scanning_prospection@ perigosa /=COB= que tem esses tipo de cobra /=COM= né //AUX=

O exemplo 38 representa duas repetições em sequência: “tem tem” e “mata mata”, rodeados por *scannings*. A anotação se dá identificando os fenômenos associados aos trechos repetidos. Infelizmente, a reflexão a respeito das diferenças entre *scanning* e *repetition* não se encontrava avançada o suficiente na época das anotações. Futuramente, vale a pena revisar se muitos dos casos anotados como *repetitions* não se tratam, na realidade, de *scannings*, principalmente quando há quebra prosódica entre os itens repetidos.

4.1.1.4. *Abandonment* (abandono)

O abandono ocorre quando o falante suspende definitivamente a sequência linguística que estava em construção, sem voltar a ela ou sem tentar expressar novamente aquele conteúdo informativo. São muito frequentes em interações conversacionais, em que mais de dois falantes disputam o turno e acabam induzidos à auto-interrupção.

O exemplo 39 mostra um caso de abandono que ocorre pela demanda do turno de REN por parte da interlocutora FLA. Antes que REN encerre seu programa informacional, FLA começa sua sequência linguística, com algumas de suas palavras se sobrepondo ao turno de REN (conforme a medida de tempo dos alinhamentos). Diante disso, REN abandona seu enunciado e não volta a construí-lo na sequência.

Exemplo 39 bfamd101 C-ORAL-BRASIL I 

*REN: tem que olhar a lista aí /=COB= <das> +=UNC=@abandonment_external interrupt@
00:00:04.880 - 00:00:07.321

*FLA: <eu lembro do> dia da festa /=COB= que a gente foi comprar bebida ali //COM=

00:00:06.891 - 00:00:10.210

*REN: hhh que que tem //COM=

00:00:10.210 - 00:00:12.096

O abandono também ocorre sem que o interlocutor intervenha no turno. O exemplo 40 mostra um caso em que, durante a produção de um enunciado, a própria falante opta por não continuar seu programa informacional sem que haja demanda do turno por parte de outras pessoas.

Exemplo 40 *bfamd101* C-ORAL-BRASIL I 

*FLA: esse também vai &sa +=UNC=@abandonment_program alteration@

*FLA: ah //COM=

*FLA: não //COM=


4.1.1.5. *Filled pause* (pausa preenchida)

A pausa preenchida é caracterizada pela produção de segmentos, geralmente vocálicos e/ou nasais, que sinalizam ao interlocutor que o falante tem a intenção de continuar o enunciado. Ela pode ocorrer logo antes do início do padrão informacional ou em seu interior. O exemplo 41 mostra duas pausas preenchidas que ocorrem em sequência, indicando que o falante estava planejando a unidade informacional seguinte. Nesse exemplo, especificamente, não há quebra prosódica entre uma pausa preenchida e outra, embora a quebra exista entre a segunda pausa e a unidade ilocucionária.

Exemplo 41 *bfamcv01* C-ORAL-BRASIL I 

*GIL: assim /=INT= &he @filled pause_prospection@ &he /=TMT=@filled pause_prospection@ quadra coberta //COM=

As pausas preenchidas também podem ocorrer como consequências de outros fenômenos de disfluência. No exemplo 42, a falante desmembra em uma *scanning* a unidade de COM em duas unidades entonacionais: “a gente acabava” e “cabava rindo”. Essa primeira disfluência parece condicionar a ocorrência da pausa preenchida, que fica entre os dois conteúdos locutivos que integram o Comentário.

Exemplo 42 *bfamnn02* C-ORAL-BRASIL I 


*DFL: <então> brincava /=COB= e /=DCT= com isso /=TOP= a gente acabava [/1]=SCA=@scanning_prospection@ &he /=TMT=@filled pause_prospection@ cabava rindo //COM=

Um estudo realizado ao longo desta dissertação propõe uma análise fonética dos segmentos produzidos em pausas preenchidas do português brasileiro e do italiano. Os corpora utilizados são o C-ORAL-BRASIL I e o C-ORAL-ROM, que são comparáveis entre si. Os resultados estão apresentados na forma de um artigo curto no Apêndice A. Em linhas gerais, é possível notar uma diferença convencional nos segmentos presentes nas pausas preenchidas dessas duas línguas. Enquanto a nasalidade não faz parte das pausas do português brasileiro, ela é amplamente atestada naquelas do italiano. A hipótese para explicar essa diferença reside no fato de que as particularidades fonológicas das línguas condicionam os segmentos que ocorrem em pausas preenchidas. O português, que atribui distintividade fonêmica à nasalidade, não seleciona esse traço fonético nas pausas; o italiano, que não possui nasalidade em seu inventário fonêmico, possibilita que as pausas sejam preenchidas por vogais nasais. Os detalhes desse trabalho podem ser melhor compreendidos no Apêndice A desta dissertação.

4.1.1.6. *Change of textual planning* (mudança de planejamento textual)


A mudança de planejamento informacional ocorre quando o falante reelabora o modo de apresentação do conteúdo locutivo da unidade informacional. Seu foco é apenas a reelaboração textual, sem alteração no programa informacional.

O exemplo 43 mostra uma mudança de planejamento textual (*change of textual planning*) que ocorre por conta da incerteza do falante em relação à data do acontecimento que está sendo narrado. A mudança ocorre na unidade de Comentário, na qual há a troca de “mil-novecentos-e-sessenta-e” por “no Natal de sessenta-e-dois”. A unidade informacional que estava sendo construída na primeira tentativa é a mesma daquela realizada após a mudança de planejamento. Nessa última, a mudança textual torna o conteúdo semântico mais específico, incluindo o período do ano em que se deu o evento (no Natal). Não se trata, portanto, de um reparo, mas da escolha de um outro modo de apresentação do conteúdo. Da unidade anterior, salva-se apenas o verbo de ligação “foi”, que estabelece composicionalidade sintática com a unidade que veicula a força ilocucionária: “foi no Natal de sessenta e dois”. Nossa análise permite dizer que a pausa preenchida também foi inserida como consequência estrutural da mudança de planejamento.

Exemplo 43 *bfamnn03* C-ORAL-BRASIL I 


*ALO: o negócio /^{=TOP=} foi mil-novecentos-e-sessenta-e [/1]^{=SCA=@change of textual planning_uncertainty@} &he /^{=TMT=@filled pause_uncertainty@} no Natal de sessenta-e-dois //^{=COM=}

No exemplo 44, a mudança de planejamento textual também ocorre no Comentário. Nesse caso, nenhuma palavra da unidade cujo planejamento foi alterado estabelece composicionalidade com a unidade que veicula a força ilocucionária. Esse é um dos casos de potencial divergência nas anotações entre os dois corpora italianos (*Corpus di Italiano Parlato* e C-ORAL-ROM) e o C-ORAL-BRASIL, pois se trata de um caso em que nenhuma palavra da primeira unidade entonacional é aproveitada na unidade seguinte. Enquanto os italianos tenderiam a anotá-lo como *retracting* sem duplicação ou ambiguidade *retracting-interrompido* ([//]) (cf. 2.2.3.1 e 2.2.3.2), o C-ORAL-BRASIL anota como um caso de *retracting* simples ([/n]) (cf. 2.2.3.3).

Exemplo 44 *bpubd101* C-ORAL-BRASIL I 

*PAU: tem que ser [/3]=EMP=@change of textual planning_program alteration@ o assunto é espontâneo //COM=

O exemplo 45 conta com um caso emblemático em que uma estrutura comparativa inicialmente evocada pela falante é substituída na mudança de planejamento. É provável que TER realizasse uma unidade de Comentário da seguinte forma: “a mãe da Fafica é mais nova que Fulana de Tal”. Em determinado momento da realização, a falante se dá conta de que essa ilocução seria melhor estruturada com outro texto - que não o comparativo. Aproveita-se então do sujeito sintático “a mãe da Fafica”, que se torna composicional com a nova estrutura “é da idade da Jael”.

Exemplo 45 *bfamcv02* C-ORAL-BRASIL I 

*TER: a mãe da Fafica é mais &n [/3]=SCA=@change of textual planning_improvement@ é da idade da Jael <hhh> //COM=

A mudança de planejamento textual se distingue da mudança de planejamento informacional, explicada na sequência.

4.1.1.7. *Change of informational planning* (mudança de planejamento informacional)

A mudança de planejamento informacional ocorre quando o falante realiza mudanças pragmáticas - majoritariamente informacionais - ao longo da sequência linguística. Para além de alterações textuais, estão presentes mudanças de unidade informacional e/ou classe ilocucionária.

O exemplo 46 mostra um caso em que a construção da unidade de Comentário foi interrompida por uma mudança de planejamento informacional (*change of informational planning*). Essa operação permitiu a inserção de uma unidade de Tópico, que estabelece o domínio temporal dentro do qual a ilocução será efetivamente realizada pelo Comentário. Diferentemente dos casos de mudança puramente no nível locutivo (como a mudança de planejamento textual), a mudança informacional se dá pelo fato de que o falante julga necessário acrescentar elementos pragmáticos (unidades informacionais) que auxiliem a interpretação da unidade ilocucionária.

Exemplo 46 bfamnn01 C-ORAL-BRASIL I 

*MAI: no norte de Minas /=TOP= tinha esse [/2]=EMP=@change of informational planning_program alteration@ antigamente /=TOP= tinha esse tipo de cobra tudo /=COM= né //=-AUX=

Como uma disfluência de nível pragmático, a mudança de planejamento informacional pode gerar outras disfluências no nível locutivo. Isso ocorre em 46 com a repetição do trecho “tinha esse” antes e depois do acréscimo do Tópico. Uma operação similar está presente no exemplo 47, em que, durante um discurso reportado, a falante realiza uma mudança de planejamento informacional para acrescentar também uma unidade de Tópico. Nesse caso, há a repetição do trecho “ele é igual” após o Tópico reportado.

Exemplo 47 bfamnn01 C-ORAL-BRASIL I 

*DFL: eu falei /=INT= eh /=CMM_r= pois é /=CMM_r= mas ele é igual [/4]=EMP_r=@change of informational planning_program alteration@ pa mim /=TOP_r= ele é igual a dos outros <hhh> //=-COM_r=

A mudança de planejamento do exemplo 48 envolve não apenas o planejamento informacional, mas, sobretudo, ilocucionário. Após fazer uma breve explicação a respeito da numeração de amostras de sangue no laboratório em que trabalha, a falante se dá conta de que não conhece tanto a respeito de um dos processos. Com esse contexto, ela inicia uma ilocução de pergunta, que acaba interrompida antes de sua finalização (“como é que &di”). Na iminência de finalizar a pergunta, a falante muda seu planejamento, nesse caso, ilocucionário, e produz uma afirmação (“eu nã sei direito como é que funciona esse esquema”).

Exemplo 48 bpubcv01 C-ORAL-BRASIL I 

*FLA: aí /=DCT= cada dia /=TOP= vai ser uma /=SCA=@scanning_prospection@ uma nova numeração /=COM= assim /=PAR= né //AUX=

*FLA: como é que é &di +=UNC=@change of informational planning_program alteration@

*FLA: eu nũ sei direito como é que funciona esse esquema /=COM= assim //PAR=

Em nossas análises, percebemos que casos como esse podem ser confundidos com reparos de ilocução. As diferenças entre as duas categorias (caso de fato existam) ainda parecem nebulosas na análise *corpus-driven*. Por ora, podemos argumentar que o reparo ilocucionário (como no exemplo 31) envolve o uso de uma ilocução inapropriada, que é prontamente substituída por outra, mantendo intacto o conteúdo semântico. A mudança de planejamento informacional-ilocucionária, por outro lado, parece envolver sempre alguma reelaboração locutiva, para além da reformulação pragmática. Essa proposta precisa ser melhor apreciada em análises futuras.

4.1.2. As categorias de motivação

São nove as motivações identificadas ao longo da categorização *corpus-driven*. A motivação de uma disfluência está mais intimamente ligada com a razão de ser do fenômeno do que a sua tipologia. Nesta seção, serão apresentadas e exemplificadas cada uma das motivações resultantes da classificação. Nas anotações, as motivações se encontram após o símbolo de underline (“_”), entre os dois sinais de arroba (“@”).

4.1.2.1. *Prospection* (prospecção)

Muitas vezes o falante produz uma disfluência pelo fato de que o conteúdo locutivo ou o padrão informacional (ou ambos) estão sendo planejados com alguma dificuldade. A prospecção ocorre justamente nesses casos, em que o falante está em busca do que dizer e/ou como dizer. Ela frequentemente é a causa de *scannings* e pausas preenchidas, conforme os exemplos 49 e 50.

Exemplo 49 bfamnn01 C-ORAL-BRASIL I 

*MAI: aí /=TOP= deu a noite /=TOP= e a mulher nũ podia sair pa poder /=SCA=@scanning_prospection@ ir atrás de alguém //COM=

Exemplo 50 bfamcv01 C-ORAL-BRASIL I 

*GIA: porque assim /=INT= &he /=TMT=@filled pause_prospection@ quando eu descobri lá o São Vicente /=TOP= eu adorei //=-COM=


No exemplo 49, o falante produz a unidade de Comentário em duas unidades entonacionais. É possível pensar que o planejamento textual da segunda unidade entonacional (“ir atrás de alguém”) - aquela que carrega o núcleo ilocucionário - não foi realizado junto com o planejamento da primeira unidade (“e a mulher nã podia sair pa poder”). Nesse caso, a *scanning* é motivada pela prospecção do conteúdo locutivo da segunda parte do Comentário.

A disfluência do exemplo 50 se trata de uma pausa preenchida, também motivada por prospecção. A única diferença com relação ao 49 é o fato de que, em 50, o falante está prospectando o restante do padrão informacional, i.e., o Comentário e o Tópico a ele indexado. A pausa preenchida funciona como uma marcação interacional de que o falante está planejando a próxima parte do enunciado.

4.1.2.2. *Inadequacy* (inadequação)

As disfluências também podem ocorrer por casos em que o falante percebe que o que foi dito é, na verdade, inadequado. A percepção de inadequação pode ocorrer tanto para o conteúdo locutivo quanto para a estrutura informacional. A maior parte dos reparos gramaticais parece ser motivada pela percepção de inadequação.

O exemplo 51 mostra um caso de reparo motivado pela percepção de inadequação, uma vez que a falante DFL começa a pronunciar a palavra “dem[orasse]” (provavelmente por uma espécie de repetição automática do trecho dito anteriormente pelo interlocutor LUC). Ao perceber a inadequação de significado, ela prontamente produz uma quebra não terminal, repete o advérbio de negação e acrescenta o verbo mais adequado (“voltasse”).


Exemplo 51 bfamvn02 C-ORAL-BRASIL I 

*LUC: <correio era um trem demorado> hhh //=-COM=

*DFL: que ele não &dem [/2]=SCA=@repair_inadequacy@ não voltasse //=-COM=

Embora não tenha sido acrescentado aos dados quantitativos, o exemplo 52 representa outro reparo que ocorre por percepção de inadequação e que possui como consequência uma *scanning*. A falante REG parece ter a intenção de produzir um enunciado composto por uma única unidade entonacional. No entanto, durante a produção da palavra “enfermeira”, a falante se dá conta de que deveria pronunciar o composto “enfermeira-chefe”. Até que esse outro item

lexical pudesse ser prospectado, a falante reinicia a estrutura a partir do núcleo do sintagma preposicional, alongando a preposição de maneira análoga a uma pausa preenchida. Quando o item “enfermeira-chefe” é recuperado, a estrutura segue normalmente.


Exemplo 52 *bfamnn04* *C-ORAL-BRASIL I* 

*REG: aí a &enfermei [/2]=SCA=@repair_inadequacy@ a /=SCA=@scanning_prospection@
enfermeira-chefe deixou ele entrar //COM=

Uma segunda proposta explicativa para esse reparo reside no fato de que a posição dos grupos de força (Câmara-Jr, 1970) em “enfermeira-chefe” e “enfermeira” são diferentes. O reparo se explicaria pela necessidade de adequação suprasegmental do vocábulo.

4.1.2.3. *Articulation problem* (problema articulatorio)

O problema articulatorio se dá quando o falante comete um erro de natureza articulatória, isto é, de pronúncia. É possível discutir em que medida essa motivação não seria a mesma da inadequação, uma vez que, para retificar o erro articulatorio, o falante deve percebê-lo como inadequado. A incorporação dessas duas categorias numa única não parece problemática, assim como pode ser útil mantê-las separadas. O exemplo 53 mostra um caso de problema de articulação num Comentário Ligado, em que a falante pronuncia acidentalmente “escoveu” ao invés de “escolheu”. O reparo então ocorre, e o COB é plenamente articulado na sequência.

Exemplo 53 *bfamnn02* *C-ORAL-BRASIL I* 

*DFL: aí ele escoveu [/1]=SCA=@repair_articulation problem@ escolheu direito /=COB=
e foi /=SCA= um bom advogado //COM=

4.1.2.4. *Improvement* (melhoramento)

O melhoramento ocorre quando o falante produz uma disfluência com o objetivo de melhorar aspectos da estrutura textual e/ou informacional. Diferentemente do senso de inadequação, em que o falante percebe que o que foi produzido é inapropriado e merece uma retificação, no melhoramento, a disfluência ocorre mesmo que o que tenha sido dito seja considerado correto e apropriado.

O exemplo 54 apresenta uma mudança de planejamento textual motivada por melhoramento (*improvement*). O falante estava no curso de uma unidade de Tópico que provavelmente seria realizada como “falando em povo mascarado”, até que se produz uma

quebra prosódica e se faz um acréscimo textual do trecho “e também”. Essa reelaboração não funciona como uma correção de um erro e nem acrescenta conteúdo proposicional ao trecho em questão. Ainda assim, o falante opta por realizar a mudança de planejamento como uma estratégia de melhorar a estrutura locutiva, estabelecendo explicitamente a coesão com os trechos anteriores.

Exemplo 54 *bfamcv01* *C-ORAL-BRASIL I* 


*GIL: <ô /=CNT= mas> /=DCT= voltando à questão /=TOP= falando em [/2]=EMP=@change of textual planning_improvement@ e também falando em povo mascarado /=TOP= esse povo do Galáticos é muito palha /=COB= eu acho que es nũ deviam mais participar /=COM= e <tal> //UNC=

O exemplo 45, citado na seção 4.1.1.6., conta com outro caso de mudança de planejamento textual motivada por melhoramento do texto.

4.1.2.5. *Uncertainty (incerteza)*

A incerteza ocorre quando o falante demonstra não possuir convicção ou segurança a respeito de uma informação. Na dúvida com relação a essa informação, o falante produz as disfluências.

O exemplo 55 representa duas pausas preenchidas motivadas pela incerteza dos falantes. A pausa preenchida de BAL é motivada pelo fato de ele estar em dúvida com relação ao local aonde guardar o cabo. A pausa preenchida de BEL ocorre pelo fato de ela não possuir uma resposta confiante para a pergunta de BAL.

Exemplo 55 *bfamd102* *C-ORAL-BRASIL I* 


*BAL: &he /=TMT=@filled pause_uncertainty@ onde que eu coloco esse cabo aqui //COM=

*BEL: &he /=TMT=@filled pause_uncertainty@ talvez na parte maior /=CMM= não //CMM=

4.1.2.6. *Program alteration (alteração do programa)*

A alteração do programa ocorre quando o falante produz disfluências que alteram o programa informacional ou textual que já estava sendo executado. Ainda não está claro se essa categoria deve servir como motivação guarda-chuva, que também inclui os casos de melhoramento (*improvement*).

O exemplo 56 apresenta uma mudança de planejamento motivada pela alteração no programa informacional. O falante estava prestes a concluir um subpadrão informacional com uma unidade de COB, quando resolve acrescentar a unidade de TOP (“negócio meu”).

Exemplo 56 *bfamnr03* C-ORAL-BRASIL I 

*AL0: eu falei /=INT= não /=PHA_r= e eu /=SCA_r=@scanning_unclear@ devo ter a conta ainda /=COB_r= porque eu nũ [/2]=SCA_r=@change of informational planning_program alteration@ negócio meu /=TOP_r= eu nũ jogo &f nada fora /=COB_r= tá tudo guardado //COM_r=

O exemplo 44 mostra um caso similar em que a alteração do programa também ocorre para o acréscimo da unidade de Tópico.

4.1.2.7. *External interrupt* (interrupção externa)


A interrupção externa é uma motivação frequentemente encontrada em interações com dois ou mais participantes (diálogos e conversações). As disfluências geradas por ela ocorrem em virtude de o falante suspender o fluxo da fala por conta de algum evento externo à sua vontade. Geralmente ocorre quando o interlocutor demanda o turno e começa um enunciado sobreposto àquele do falante.

O exemplo 57 mostra um caso em que GIL produz um abandono (*abandonment*) motivado pelo fato de que GIA começa sua sequência linguística. Nesse caso, a demanda pelo turno por parte de GIA se configura como um evento externo à vontade de GIL, que o faz desistir do enunciado que estava começando a construir.

Exemplo 57 *bfamcv01* C-ORAL-BRASIL I 

*GIL: <não> /=COB= e assim /=COB= esse que é o ponto //COM=
 *GIL: eles <reclamavam> +=UNC=@abandonment_external interrupt@
 *GIA: <não /=AUX_n= se bem que tinha> /=SCA=@scanning_external interrupt@ tinha uns dois que falaram assim /=INT= gente /=AUX_r= a gente perdeu no campo /=CMM_r= a gente nũ perdeu por causa disso não //CMM_r=

O exemplo 58 faz parte do C-ORAL-BGEST e seu arquivo conta com alinhamento gestual. A falante ELI realiza um monólogo sobre sua experiência como professora de inglês. Em determinado momento, entra na sala sua gata de estimação, que a faz interromper o fluxo da fala e mudar seu foco de atenção. O evento externo, nesse caso, não é a fala do interlocutor, mas a chegada do animal ao ambiente de gravação.

Exemplo 58 bgest_06 C-ORAL-BGEST 

*ELI: então /=DCT= eu pedia pro meu aluno falar do trabalho dele /=COB= em inglês obviamente /=PAR-COB= a gente tinha livro /=PAR-COB= e tudo /=PAR-COB= mas eu sempre levava /=SCA= pro trabalho /=SCA= do aluno //COM=

*ELI: que isso ia ajudá-lo +=EMP=@change of informational planning_external interrupt@


*ELI: olha quem chegou //COM=

No caso do exemplo 58, inclusive, é possível notar que a unidade gestual que estava sendo construída acabou sendo paralisada na mudança de planejamento. O gesto interrompido, no entanto, se mantém em posição de *hold* enquanto a falante faz o comentário sobre a chegada da gata. Esse caso é discutido em Albuquerque (2024), evidenciando a íntima relação entre gesto e fala, mesmo nos momentos de disfluência (Graziano & Gullberg, 2018).

4.1.2.8. Immediate response (resposta imediata)

A resposta imediata é uma motivação de disfluência que comumente é confundida com interrupção externa. Ambas podem ser geradas por elementos externos, mas possuem uma diferença pragmática crucial. Na resposta imediata, o falante suspende seu enunciado em curso para fornecer uma ilocução imediata em resposta ao evento externo. Na interrupção externa, isso não ocorre, e o falante tem seu enunciado interrompido sem fornecer uma resposta ilocucionária à demanda externa.

O exemplo 59 representa um abandono gerado por resposta imediata. A falante TER interrompe sua sequência após o enunciado de JAE. Essa interrupção parece ocorrer precisamente para que TER possa dirigir uma outra ilocução (“ô Jael”) à interlocutora. Essa nova ilocução tem a função de realizar uma espécie de advertência em face do que foi dito por JAE (“ea nũ tem nada na mão”).

Exemplo 59 bfamcv02 C-ORAL-BRASIL I 

*JAE: <ganhou não> //

*TER: <não> //


*TER: <vai ganhar /=COB= mas> +=UNC=@abandonment_immediate response@

*JAE: <ea nũ tem nada> na mão //COM=

*TER: ô Jael //COM=


O exemplo 60 mostra um caso interessante que ocorre um pouco mais à frequente nessa mesma gravação. Nesse contexto, as participantes estão aventando nomes de parentes que

poderiam dar um sofá de presente a um casal de noivos. No primeiro turno, TER realiza um introdutor locutivo, prestes a iniciar um trecho de ilocuções reportadas. A interlocutora JAE, no entanto, na tentativa de adivinhar o desfecho do enunciado reportado, aproveita o turno para dizer que a falante TER ficará responsável por presentear os noivos com o sofá. TER então interrompe temporariamente seu enunciado, fornece uma resposta à ilocução de JAE e retoma em seguida a sequência reportada que havia começado no primeiro turno.

Exemplo 60 *bfamcv02* *C-ORAL-BRASIL I* 

*TER: <aí eu> falei assim então <hhh /=INT= o &s> [/2]=EMP=@repetition_unclear@
 o &s +=UNC=@abandonment_immediate response@
 *JAE: <ocê vai dar> //COM=
 *TER: não //COM=
 *TER: o sofá vão ver /=COB_r= que ea às vezes a Rute distribui com ela
 /=COB_r= Tonita /=COB_r= <Ção> //COM_r=

O mesmo exemplo pode ser ouvido, pelo repositório de exemplos da dissertação com e sem a ilocução “não” realizada de forma imediata ao estímulo de JAE. O exemplo editado, isto é, como em 61, é absolutamente razoável do ponto de vista perceptual. Os trechos reportados, como já discutido no capítulo 2, representam uma mudança do aqui e do agora da interação. Neles, os falantes emulam vozes de outra pessoa ou deles próprios em outro momento do tempo.

Exemplo 61 *bfamcv02* *C-ORAL-BRASIL I* 

*TER: <aí eu> falei assim então <hhh /=INT= o &s> [/2]=EMP=@repair_articulation
 problem@ o &s [/2]=EMP=@repetition_unclear@ o sofá vão ver /=COB_r= que ea às vezes a Rute
 distribui com ela /=COB_r= Tonita /=COB_r= <Ção> //COM_r=

A operação de resposta imediata, nesse caso, produz uma complexa suspensão temporária da reportação - o que implica no retorno ao aqui e agora da interação - para fornecer uma ilocução no nível do enunciado. Ao fim da ilocução, a falante retoma o nível do discurso reportado. A disfluência funciona apenas como uma consequência de uma operação pragmática complexa, e sua ocorrência espelha um domínio absoluto das funções ilocucionárias e do *displacement* no enunciado reportado.

4.1.2.9. *Satisfaction* (contentamento/satisfação)

O contentamento ou satisfação ocorre quando o falante suspende o discurso por já se encontrar satisfeito com o que foi dito. Nesses casos, ele considera que não precisa dar continuidade à sequência linguística que estava sendo construída. Os casos de satisfação geralmente levam a abandonos (*abandonments*) ou a mudanças de planejamento informacional (*change of informational planning*).

O exemplo 62 representa um caso de mudança de planejamento informacional causada pelo contentamento da falante, FLA, com relação à veiculação das informações. Ela inicia fazendo uma lista de doenças que podem estar presentes em amostras de sangue. Na iminência de citar a terceira, o interlocutor BRU sinaliza que compreendeu a situação, então a falante se faz satisfeita e muda seu planejamento informacional.

Exemplo 62 bpubcv01 C-ORAL-BRASIL I 62

*FLA: só que a gente não sabe /=COB_s= se elas têm agaivé /=CMM= se elas <têm hepatite> /=CMM= se elas +=UNC=@change of informational planning_satisfaction@

*BRU: <ah /=CMM= só> //CMM=

*FLA: a gente nũ sabe nada sobre elas //COM=

4.2. Caracterização estatística da amostra

Foram encontradas 944 ocorrências de disfluências na amostra selecionada de dez textos do minicorpus do C-ORAL-BRASIL I. Todas elas foram anotadas conforme os tipos e as motivações descritas acima. Os tipos *scanning* (32,4%), *repair* (17,8%), *repetition* (13,8%) e *abandonment* (12,0%) foram os mais frequentes (figura 5). Para essa anotação, houve poucos casos (3,6%) em que não foi possível definir o tipo de fenômeno corrido (casos *unclear*).

As *scannings* foram quase duas vezes mais frequentes do que os reparos. É importante frisar que o fenômeno que estamos chamando de *scanning* não necessariamente coincidem com as tradicionais escansões da L-AcT. Exemplos como aqueles mostrados em 35 e 36, na subseção 4.1.1.2, são categorizados como *scanning*, mesmo que sejam etiquetados como *retractings* EMP.

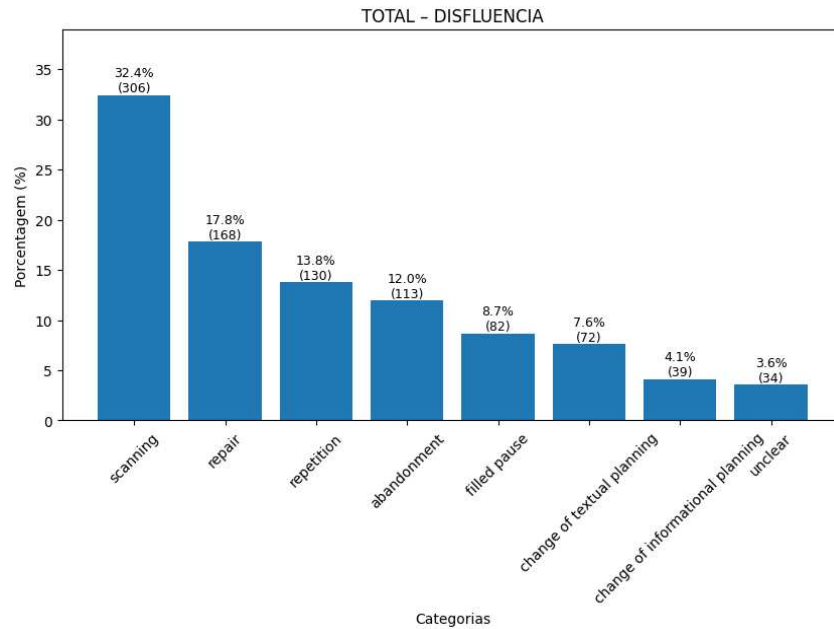


Figura 5. Total de tipos de disfluências anotados.

As motivações estão representadas na figura 6. Aquelas mais frequentes foram *prospection* (32,9%), *external interrupt* (10,5%) e *program alteration* (9,0%). Diferentemente das anotações de tipos de disfluências, a categoria *unclear* é muito mais presente para as motivações, tendo sido usada em 21,2% dos casos. Isso se deve ao fato de que o trabalho do anotador em decidir por uma motivação exige mais inferências de natureza probabilística do que a decisão pelo tipo de disfluência. Em muitos casos, por exemplo, não fica claro se o falante produz a disfluência pois tentou melhorar o texto do enunciado - e, portanto, motivado por *improvement* -, ou se simplesmente resolveu alterar o programa informacional, como em uma operação de *program alteration*. As categorias de *immediate response* (2,2%), *uncertainty* (1,6%) e *satisfaction* (1,0%) foram as menos utilizadas nos dez textos anotados.

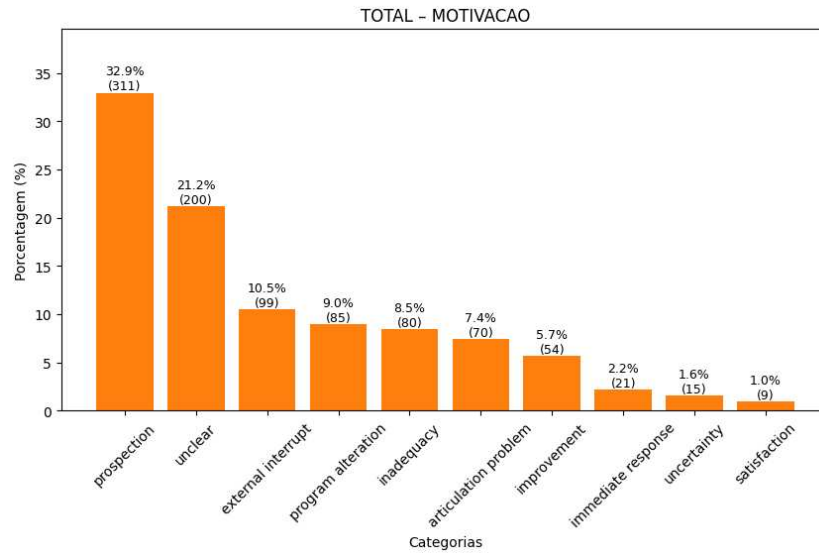


Figura 6. Total de motivações anotadas.

Foram atestadas 46 combinações entre disfluências e motivações, dentro de um rol de 80 possibilidades (assumindo que qualquer disfluência pode se associar a qualquer motivação). As frequências são bastante diversas, e apenas 7 representam mais de 5% do total de combinações. O arranjo mais frequente foi *scanning-prospection* (20,8%), seguida por *scanning-unclear* (9,2%), *filled pause-prospection* (8,1%) e *repair-inadequacy* (8,1%) (figura 7). As motivações que não puderam ser definidas (*unclear*) se associaram consistentemente a *scanning* (87 casos) e a *repetition* (71 casos).

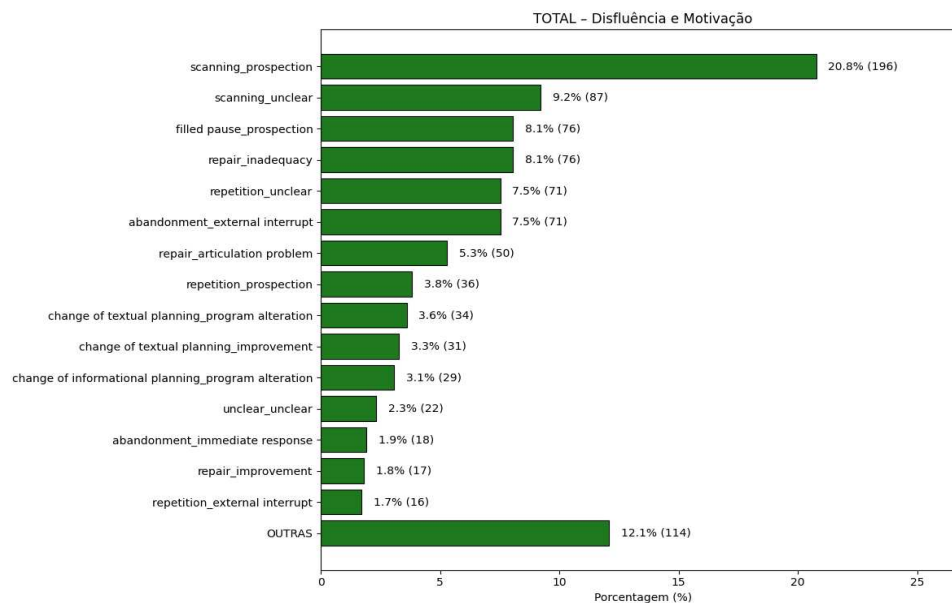


Figura 7. Valor total e relativo de associações entre disfluências e motivações.

4.3. As disfluências e suas motivações

A figura 8 representa oito gráficos com a proporção das motivações (eixo x) que ocorreram com cada uma das categorias de disfluência. O abandono (*abandonment*) tende a ocorrer motivado por interrupção externa (*external interrupt*) (62,8% dos casos), enquanto que a mudança de planejamento informacional (*change of informational planning*) ocorre geralmente por conta de alteração do programa (*program alteration*) (74,4%). De forma parecida, a mudança de planejamento textual (*change of textual planning*) pode ser motivada tanto por alteração do programa (*program alteration*) (47,2%) quanto por melhoramento (*improvement*) (43,1%). Esse resultado, especificamente, funciona como um argumento a favor da união entre *program alteration* e *improvement* numa única categoria de motivação. As mudanças de planejamento (*change of informational planning e change of textual planning*) raramente ocorrem em virtude de outras motivações.

A maioria das pausas preenchidas (*filled pauses*) ocorre por conta de prospecção (*prospection*) (92,7%), indicando que há forte associação entre essa motivação e aqueles fenômenos. O reparo (*repair*) ocorre por um conjunto mais distribuído de motivações: 45,2% dos casos são por inadequação (*inadequacy*), enquanto que 29,8%, por problema de articulação (*articulation problem*)¹⁶. Cerca de 10% deles ocorrem por melhoramento (*improvement*).

As repetições são a categoria com a maior quantidade de motivações *unclear* (54,6%), seguidas de *prospection* (27,7%) e *external interrupt* (12,3%). As *scannings* ocorrem majoritariamente motivadas por *prospection* (64,1%). Uma grande parte delas, no entanto, foi anotada como motivação *unclear* (28,4%).

A maior parte dos casos de disfluências *unclear* foram também anotadas com motivação *unclear* (64,7%). Os casos de disfluências anotadas como *unclear* somam 34 ocorrências e geralmente representam instâncias em que não foi possível determinar com clareza os fenômenos ocorridos - geralmente por conta de momentos em que o trecho se apresentava inaudível ou por pouco material segmental realizado durante sobreposição. Problemas de articulação (*articulation problem*) são o segundo fator mais representativo das disfluências *unclear* (23,5%).

¹⁶ É possível pensar em argumentos a favor da união dessas duas motivações em uma única categoria, uma vez que os erros de produção (*articulation problem*) só são reparados quando percebidos como inadequados.

Associações de Disfluências com Motivações (%)

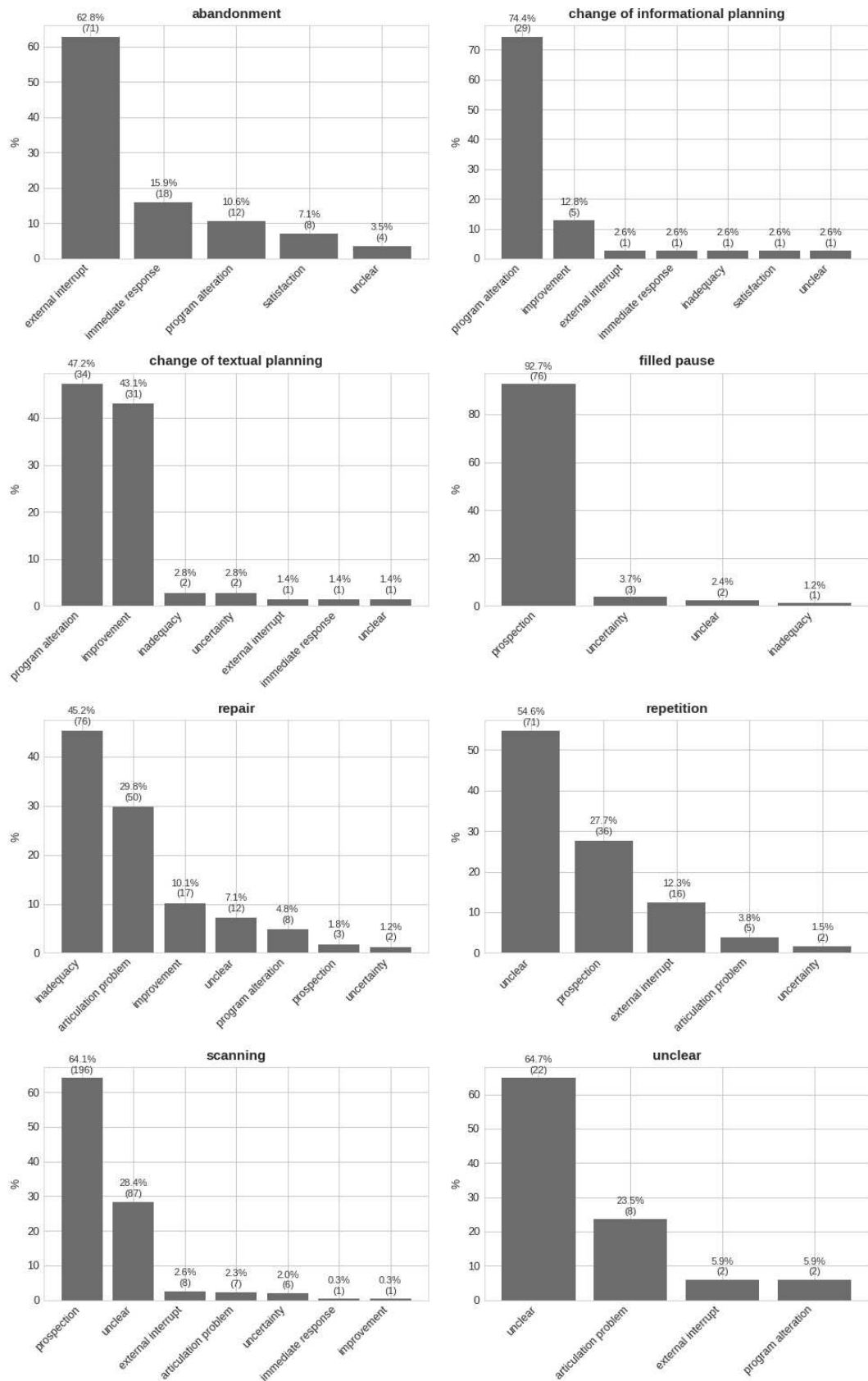


Figura 8. Frequência das motivações (%) associadas a cada uma das disfluências.

A partir da figura 7, é possível observar disfluências que são típicas de algumas motivações. As interrupções externas parecem condicionar majoritariamente os abandonos, enquanto que as alterações de programa geram mudanças de planejamento informacional ou textual. As prospecções, por sua vez, podem gerar pausas preenchidas ou *scannings*. Essas generalizações são puramente baseadas na frequência das associações e não possuem o objetivo de enrijecer as possibilidades de classificação. Elas também indicam a robustez do processo de anotação, que, embora não seja acompanhado de validação estatística prevista em trabalhos de corpora padrão-ouro, representa consistência. Em casos de anotações pouco confiáveis, se esperam resultados próximos da aleatoriedade.

4.4. As disfluências e a arquitetura da amostra

A figura 9 representa a proporção de cada um dos tipos de disfluências em função das tipologias (monólogos, diálogos e conversações). O número de palavras nas três tipologias é diferente, conforme a tabela 6 (cap. 3). Portanto, a frequência das disfluências foi normalizada para cada 1000 (um mil) palavras, operação que faz com que suas proporções se tornem comparáveis.

Os monólogos (barras azuis) possuem maior quantidade de *scanning*, *repair*, *repetition* e *filled pause*. Por outro lado, essa tipologia conta com poucos casos de *abandonments*, como é de se esperar, já que esse fenômeno é causado principalmente por interrupções externas (figura 8). A grande quantidade de *scanning*, *repair* e *repetitions* nos monólogos pode ser reflexo da menor competição do turno conversacional. Essa propriedade parece condicionar disfluências motivadas majoritariamente por processos de prospecção. Nessas gravações, os falantes geralmente narram acontecimentos do passado e possuem maior liberdade para a criação de padrões informacionais mais complexos. O mesmo não ocorre nos diálogos e nas conversações, que são tipologias mais acionais, ou seja, com maior variação de classes ilocucionárias, e menos elaboradas informacionalmente.

Os diferentes padrões de disfluência podem então ser divididos em dois grupos com relação às tipologias. O primeiro grupo é o dos monólogos, com mais *scannings*, reparos e pausas preenchidas. O segundo grupo é o dos diálogos e das conversações, marcadamente com maior quantidade de abandonos. As disfluências mais típicas dos monólogos parecem ser condicionadas pela maior complexidade dos padrões informacionais, enquanto que aquelas dos diálogos e das conversações são mais suscetíveis à natureza interacional dessas duas tipologias.

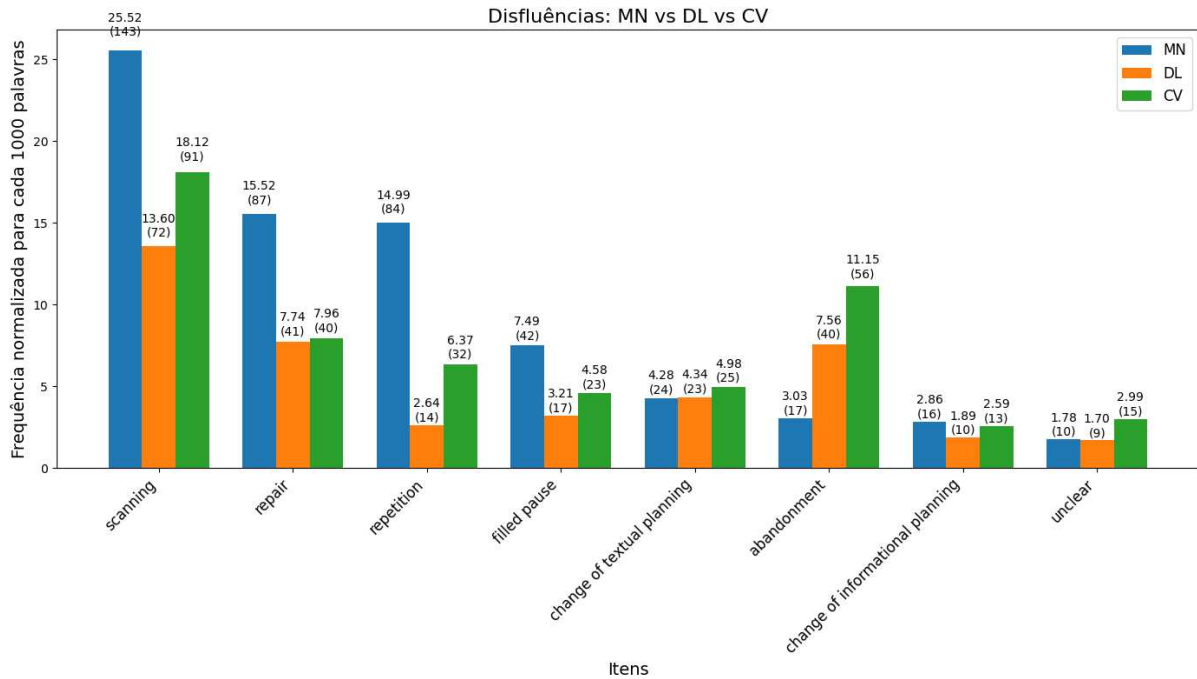


Figura 9. Disfluências encontradas nas tipologias monológica (azul), dialógica (laranja) e conversacional (verde). Para viabilizar a comparação, os dados foram normalizados para uma quantidade de 1000 (um mil) palavras; ex.: para cada 1000 palavras em monólogos, há 25,52 *scannings*. O valor entre parênteses indica o número bruto de ocorrências.

Ainda sobre a figura 9, a proporção de mudanças de planejamento textual e informacional também são comparáveis entre as três tipologias, indicando que o número de participantes em interação não condiciona esses tipos de disfluência.

A figura 10, abaixo, representa a proporção de disfluências nos dois contextos documentados: privado (azul) e público (laranja). O número de palavras da amostra utilizada também não é balanceado segundo os dois contextos. Da mesma forma que o anterior, o gráfico em questão conta com a normalização para cada 1000 palavras, o que possibilita a nossa comparação. Exceto pelo tipo *scanning*, que é mais expressivo em contexto público, todas as disfluências se mostraram mais frequentes no contexto privado. A alta frequência de *scannings* no ambiente público pode se explicar pela importância que o falante confere ao planejamento da sequência linguística. Isso condiciona muitos momentos de prospecção (principal motivação de *scannings*). Os outros tipos de disfluência são consistentemente mais frequentes em contextos privados.

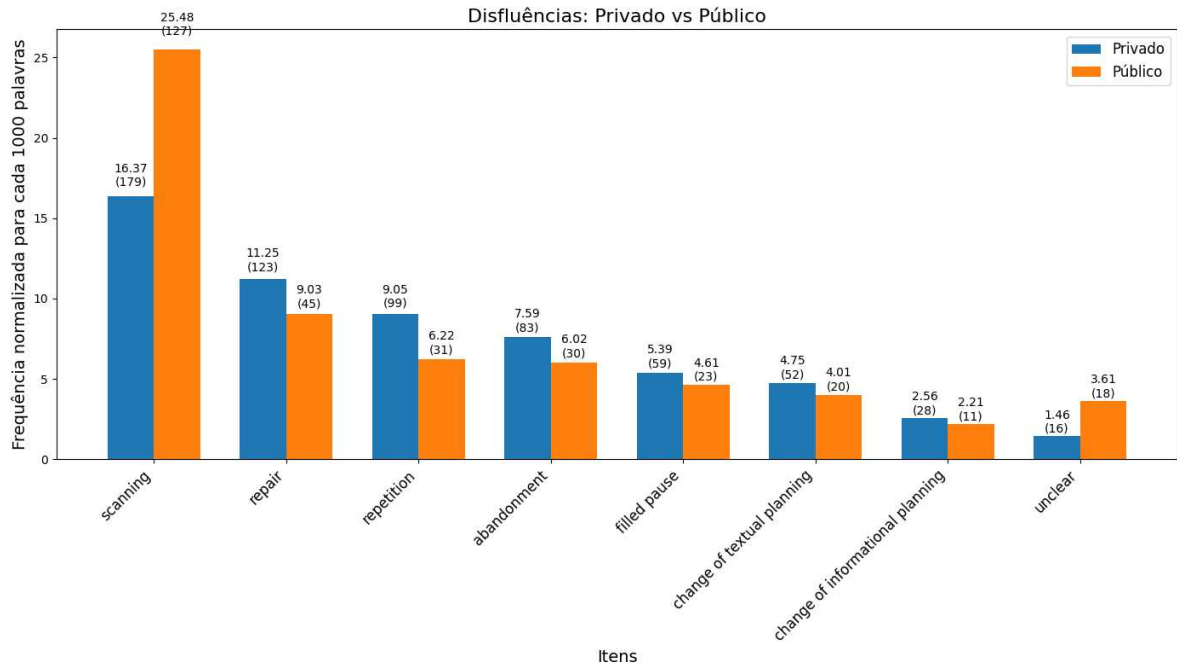


Figura 10. Disfluências encontradas nos contextos privado (azul) e público (laranja). Para viabilizar a comparação, os dados foram normalizados para uma quantidade de 1000 (um mil) palavras; ex.: para cada 1000 palavras no contexto privado, há 16,37 scannings. O valor entre parênteses indica o número bruto de ocorrências.

Dentre os dois aspectos considerados da arquitetura do minicorpus, aquele que mais parece influenciar a ocorrência de tipos específicos de disfluência é a tipologia (monológica, dialógica ou conversacional). Por outro lado, o contexto (se público ou privado) se mostra mais determinante na frequência das disfluências. Em ambos os casos, as configurações extralinguísticas geram dinâmicas comunicativas distintas que motivam de maneira diferente a ocorrência dos fenômenos estudados.

4.5. As disfluências e as unidades informacionais

Com o corpus etiquetado informacionalmente segundo as unidades previstas pela *Language into Act Theory*, foi possível quantificar a frequência das unidades informacionais realizadas com algum tipo de disfluência. Pela natureza acional da fala espontânea, as unidades informacionais não possuem as mesmas taxas de ocorrência. Unidades ilocucionárias, especialmente Comentários (COM) e Comentários Ligados (COB) ocorrem com maior frequência do que Apêndices de Tópico (APT), por exemplo, que possuem função mais específica. A tabela 7 mostra os valores brutos de cada uma das unidades etiquetadas na amostra de textos utilizada neste trabalho.

Tabela 7. Unidades informacionais encontradas na amostra e suas frequências brutas.

Unidade	Frequência
COM	2125
COB	529
CMM	634
TOP	343
APT	14
APC	68
INT	110
PAR	107
TPL	15
AUX	318
CNT	39
ALL	32
INP	21
PHA	2
DCT	98
UNC	187

A ampla diferença de frequência exige certo cuidado para normalização e interpretação dos dados de estrutura informacional. *A priori*, quaisquer fenômenos de interesse da fala estariam sujeitos a ocorrerem com mais frequência nas unidades ilocucionárias, pois são aquelas de maior quantidade bruta. Para evitarmos esse viés e podermos verificar quais são as unidades informacionais que mais ocorrem com disfluências, foi necessário estabelecer uma normalização da seguinte forma: para cada uma das unidades, calculamos quantas ocorrem com algum tipo de disfluência e quantas ocorreram sem disfluência. O resultado é apresentado em porcentagem.

A figura 11 representa a proporção de cada unidade com disfluência anotadas na amostra. É notável que 46,7% (50/107) dos Parentéticos ocorreram com alguma disfluência, contra 33,0% (175/529) de COBs e 31,4% (108/343) de TOPs. Embora sejam as unidades informacionais mais frequentes, apenas 15,8% dos Comentários foram realizados com disfluência. Do mesmo modo, apenas 7,7% dos CMMs identificados ocorrem com esses fenômenos. As unidades de auxílio dialógico, também de frequência razoavelmente elevada,

foram aquelas que atraíram a menor quantidade de disfluências, já que apenas 1,9% dessas unidades foram realizadas com algum fenômeno de fragmentação.

Embora também estejam representados como grandes focos de disfluências na figura 11, os valores para Lista de Tópicos (TPL) e para os Apêndices de Tópico (APT) não são confiáveis, uma vez que a quantidade bruta dessas unidades é muito pequena, com 15 e 14 ocorrências, respectivamente. Além disso, os valores das etiquetas de EMP e UNC foram removidos do gráfico, pois se tratam de unidades vazias que sempre representam algum tipo de disfluência. A etiqueta de UNC, particularmente, foi utilizada no minicorpus para representar a última unidade cuja construção foi abandonada num enunciado interrompido (“+”)

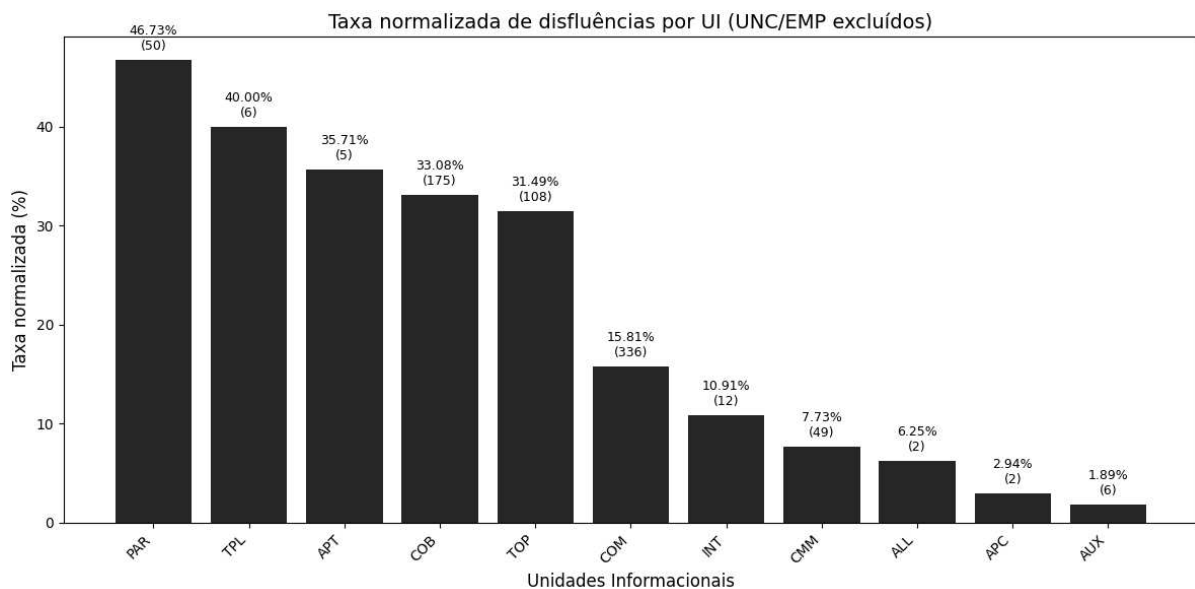



Figura 11. Frequência relativa das unidades informacionais com disfluências. A frequência foi normalizada para cada 100 unidades informacionais. Exemplo: 46% dos Parentéticos foram realizados com algum tipo de disfluência.

O exemplo 63 representa um caso em que duas disfluências ocorreram em sequência no nível Parentético. O falante inicia um o enunciado no nível reportado, mas se dá conta de que seria importante deixar claro à interlocutora quem é o indivíduo cuja fala está sendo reproduzida durante a reportação. Ele então realiza uma operação complexa, que consiste em sair do nível do discurso reportado e iniciar o nível Parentético, sem passar pelo aqui e agora do enunciado. Dentro do Parentético, ele tem dificuldade em prospectar o nome de Marcelo Crivella; quando finalmente se recorda do nome, ele retoma o reportado repetindo o conteúdo locutivo que já

havia sido apresentado por meio da unidade de Tópico. As operações do Parentético parecem ser complexas o suficiente para que mais disfluências sejam condicionadas nessas unidades.

Exemplo 63 *bfamd102* C-ORAL-BRASIL I 

*BRU: a minha igreja /=TOP_r= o [/1]=EMP=@change of textual planning_improvement@ isso
aquele /=SCA=@scanning_prospection@ Marcelo Crivella /=PAR= <aquele bispo> lá /=PAR(1)= a
minha igreja não aceita homossexuais //COM_r=

Os gráficos da figura 12 representam a proporção de cada unidade informacional realizada com alguma disfluência. É possível notar que diferentes tipos de disfluência são suscetíveis a diferentes unidades informacionais. As unidades marcadas como UNC parecem ser as que mais comportam mudanças de planejamento textual (13,4%) e informacional (13,9%). Isso significa que, do total de unidades anotadas como UNC (187), 51 foram etiquetadas como algum tipo de mudança de planejamento. Um total de 15% (16/107) das unidades de Parentético ocorrem com reparos. As *scannings*, por sua vez, se distribuem de forma mais equilibrada em unidades textuais: elas ocorrem em 13,4% dos Comentários Ligados (COBs), que são ilocuções tipicamente processuais; uma boa parte dos Parentéticos (12,4%) também conta com *scannings*. As repetições são uma categoria com grande distribuição entre diferentes unidades: os Parentéticos (8,4%), os UNC (6,4%) e os Tópicos (5,0%) são as unidades mais representativas proporcionalmente desse tipo de disfluência. Um total de 49,7% (93/187) dos UNCs são associados a abandonos de enunciado. As pausas preenchidas também parecem ter propriedades processuais, já que 4,9% dos COBs ocorrem com esses fenômenos.

É interessante notar que os Comentários Múltiplos são muito pouco frequentes (quando são atestados) para todos os tipos de disfluência. Do mesmo modo, as unidades dialógicas, que não acrescentam conteúdo semântico e nem são posicionais sintaticamente com as outras partes da sequência linguística. Para a geração da figura 12, foram excluídas as unidades com frequência menor que 40 ocorrências na amostra (APT, TPL, CNT, ALL, INP e PHA). Essa decisão foi tomada com o objetivo de não se criarem vieses percentuais baseados em unidades de baixa frequência.

Frequência relativa das UI segundo os tipos de disfluência

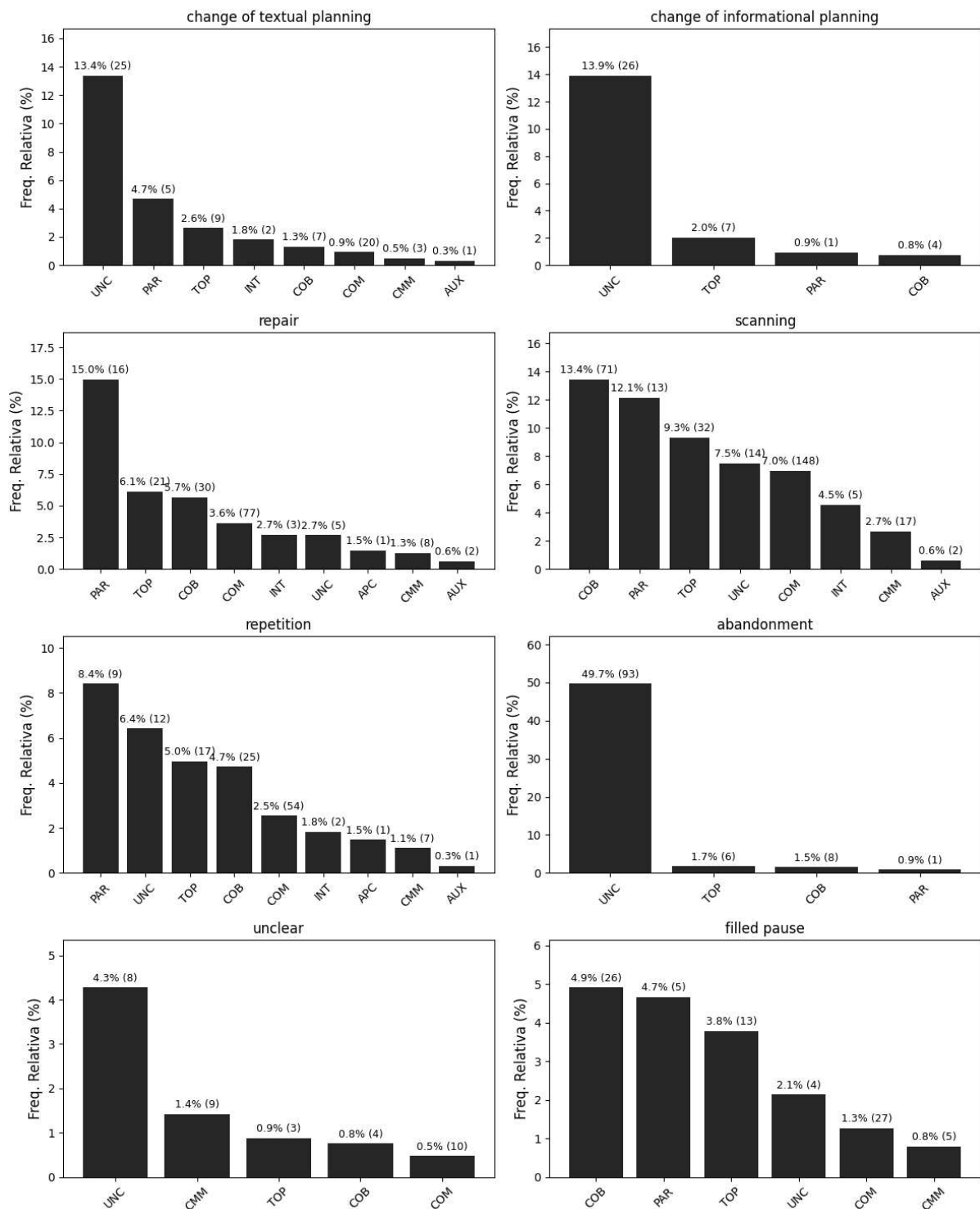


Figura 12. Frequência relativa de unidades informacionais nas quais as disfluências ocorreram. Os valores (relativos e brutos) indicam a quantidade de disfluências que ocorreram nas unidades especificadas. A frequência foi normalizada para cada 100 unidades informacionais. Exemplo: em 15% dos Parentéticos (16/107), ocorrem reparos.

Como análise complementar, foram identificadas também as unidades informacionais mais frequentes para cada um dos tipos de motivação (figura 13). As unidades em que mais ocorreram, proporcionalmente, disfluências motivadas por melhoramento (*improvement*) foram o PAR (2,8%) e o COB (2,6%). Uma parte das unidades UNC (3,7%) também tiveram disfluências classificadas como melhoramento. Nesse último caso, consideramos que o falante realizou a disfluência ao final da sequência linguística (sem terminar a última unidade) motivado por *improvement*.

Ainda sobre os gráficos da figura 13, a motivação de alteração do programa (*program alteration*) também foi bastante frequente em final de enunciado interrompido, ocorrendo em 26,2% dos UNCs. Uma proporção de 6,5% dos Parentéticos ocorreu com disfluências motivadas por problemas de articulação (*articulation problem*). Essa medida é relevante, já que o nível parentético tende a ter taxa de elocução média mais alta do que o nível do enunciado. Os casos de motivação *unclear* se distribuíram ao longo de muitas unidades informacionais, com destaque para os PAR (12,1%). O mesmo pode ser dito de disfluências motivadas por *prospection*, com PAR assumindo a liderança relativa (17,8%), seguida de COBs (15,1%) e TOPs (12,5%).

Foram anotados apenas nove casos de *satisfaction* (contentamento/satisfação). Oito deles ocorreram em final de enunciado interrompido, representando 4,3% dos UNCs. Apenas um caso foi documentado em COB. Algo similar é observado com as *immediate responses* (respostas imediatas), que somaram apenas 21 ocorrências - a maior parte delas em final de enunciado interrompido (UNC). Embora também acumulem baixa frequência bruta, os casos de *uncertainty* (incertezas) se distribuíram em TOP, UNC e unidades ilocucionárias. As *inadequacies* (inadequações) foram relativamente frequentes em unidades textuais, com destaque para os TOPs, representando 3,2% dos casos das ocorrências dessa unidade.

Assim como observado para os tipos de disfluência, as diferentes motivações seguem padrões informacionais consistentes. Boa parte delas parece estar relacionada a disfluências que geralmente são anotadas em posição final de sequências interrompidas (*improvement, program alteration, external interrupt, satisfaction e immediate response*, especificamente). As motivações *articulation problem, prospection, inadequacy e uncertainty* parecem ser típicas do encadeamento processual, no interior do enunciado.

Frequência relativa das UI segundo as motivações de disfluência

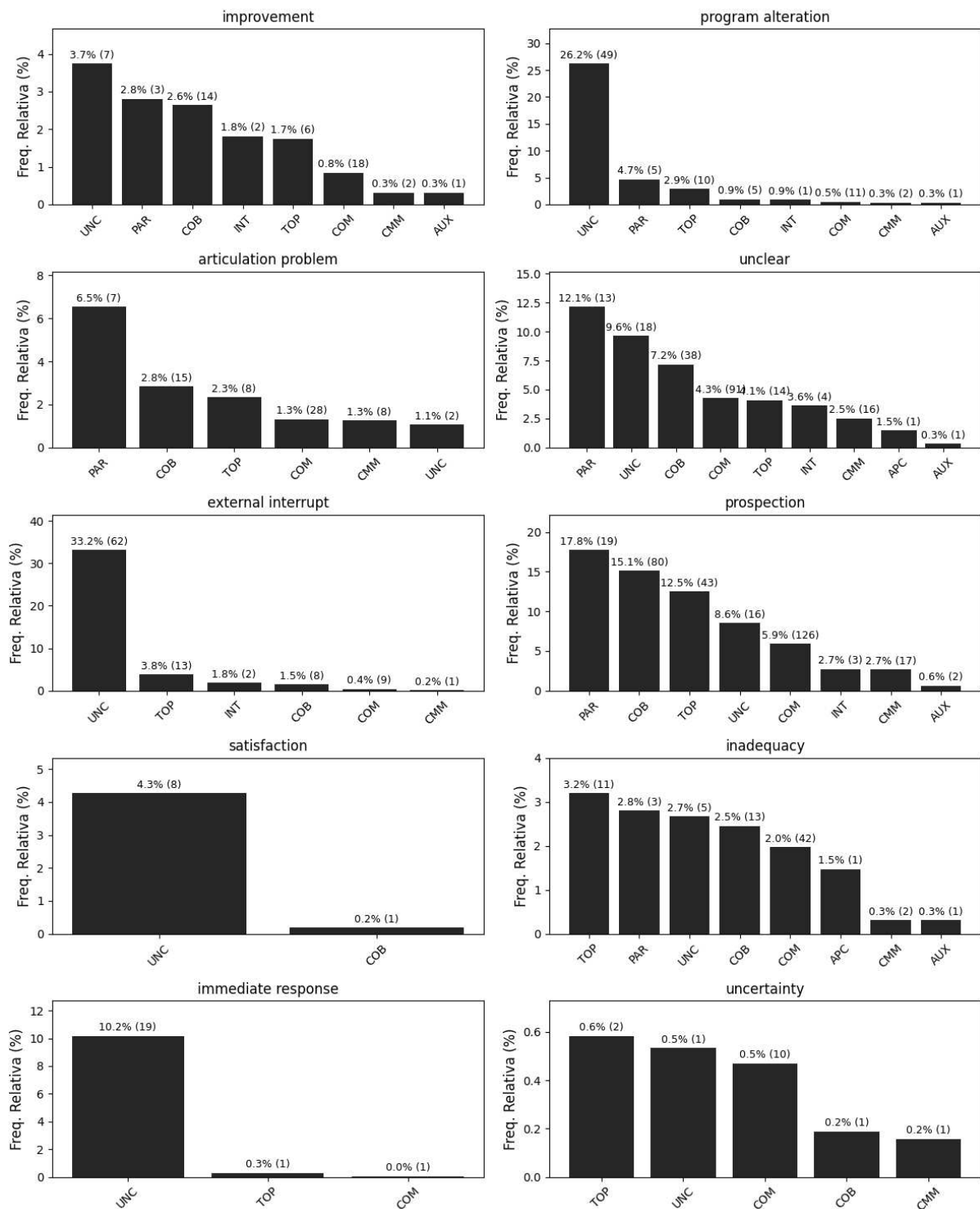


Figura 13. Frequência relativa de unidades informacionais nas quais as motivações ocorreram. Os valores (relativos e brutos) indicam a quantidade de unidades informacionais associadas à motivação do gráfico. A frequência foi normalizada para cada 100 unidades

informacionais. Exemplo: em 6,5% dos Parentéticos (7/107), ocorrem com disfluências cuja motivação foi classificada como articulation problem.

De forma geral, das figuras 11, 12 e 13, pode-se concluir que as unidades que mais atraem variados fenômenos de disfluência e suas motivações são aquelas ilocucionárias (COM, COB e CMM) e algumas textuais (PAR, TOP e INT). Nas figuras 14 e 15, isolamos apenas a frequência dos tipos de disfluências em unidades ilocucionárias e de suas motivações, respectivamente. Nesses gráficos, cada uma das barras representa a totalidade dos fenômenos indicados no eixo x. Optamos por fazer uma separação entre as unidades de interesse (as ilocucionárias), as unidades que ocorrem em interrupções (UNC), que são muito frequentes para determinados fenômenos de disfluência, e as outras unidades informacionais.

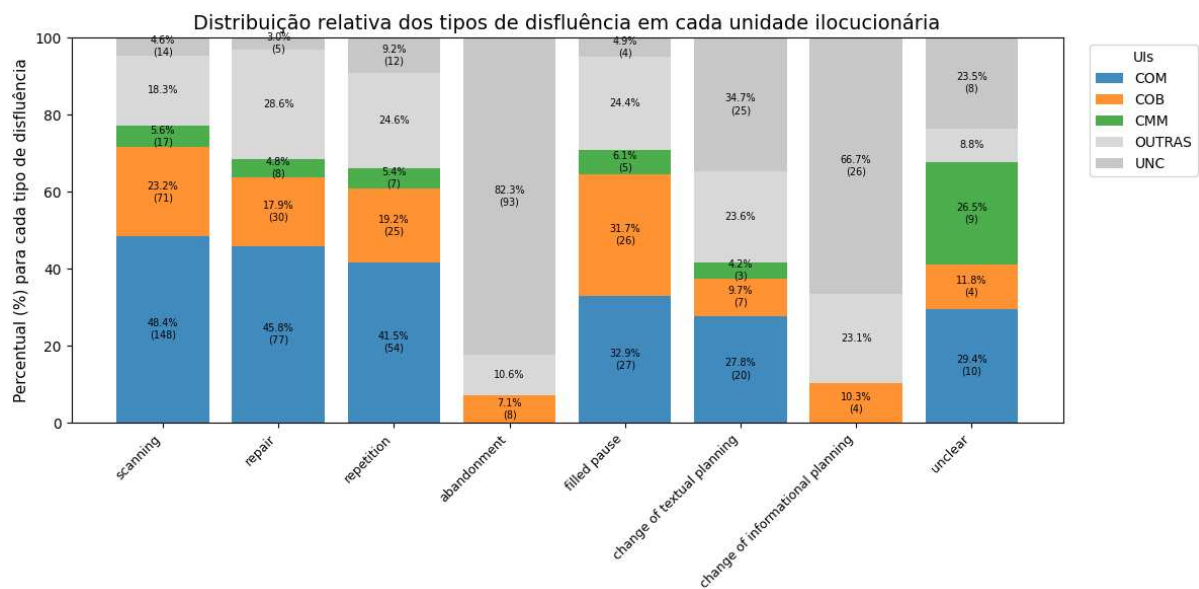


Figura 14. Frequência relativa de disfluências em cada uma das unidades ilocucionárias. Exemplo: 48,4% das *scannings* ocorrem em unidades de COM.

Com base na figura 14, é possível perceber que *scannings*, *repair*, *repetitions* e *filled pauses* ocorrem com mais frequência em unidades ilocucionárias. Os casos de *abandonment* e *change of informational planning* ocorrem mais frequentemente em final de enunciado interrompido e em outras unidades. A proporção é razoavelmente equilibrada em *change of textual planning*. Os casos de disfluência *unclear* são mais frequentes em ilocuições.

É importante ter em mente que as unidades ilocucionárias também são as mais frequentes da fala. A normalização do número de unidades - que garante a comparabilidade

entre todas elas - só pode ser visualizada nas figuras 12 e 13. Ainda assim, mesmo que as outras unidades (e o UNC) não ocorram com maior frequência, os abandonos e as mudanças de planejamento parecem ser as disfluências que melhor caracterizam os chamados enunciados interrompidos. Isso pode ser visto principalmente pela grande quantidade de UNC's registrados nesses casos.

A figura 15 representa, para as mesmas unidades ilocucionárias, a proporção de tipos de motivação. Os tipos *external interrupt*, *program alteration* e *immediate response* - que são aqueles que se associam com mais frequência aos *abandonments* e às *change of plannings* - são os menos frequentes em unidades ilocucionárias.

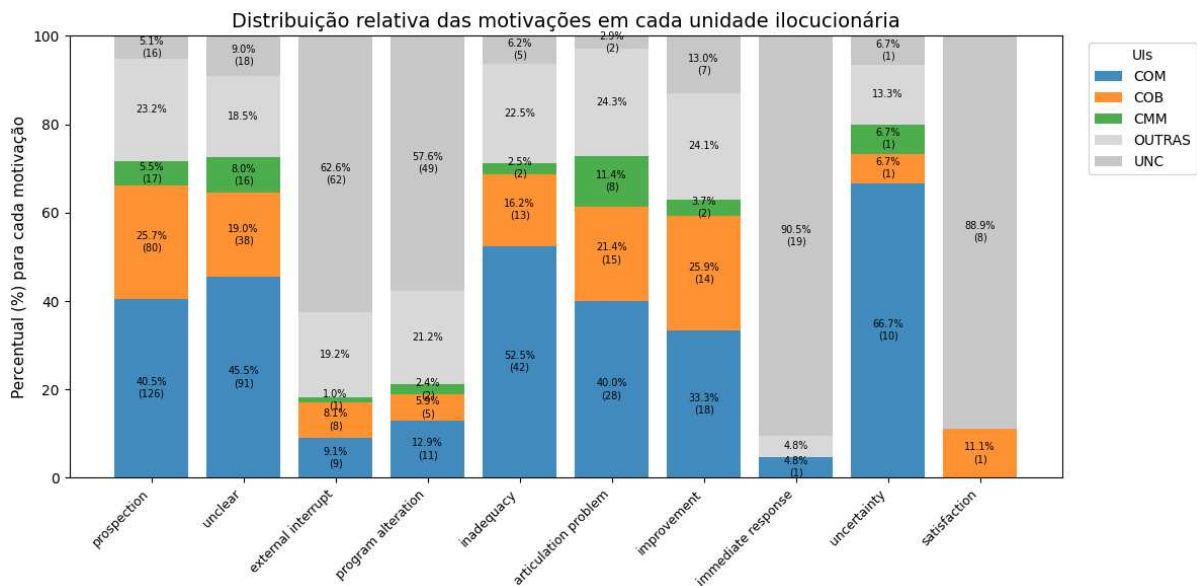


Figura 15. Frequência relativa de motivações de disfluências em cada uma das unidades ilocucionárias. Exemplo: 40,5% das *prospection*s ocorrem em unidades de COM.

4.6. Uma proposta *corpus-driven* para o tratamento das disfluências

Nossa proposta parte do pressuposto de que as disfluências permeiam todos os âmbitos linguísticos, inclusive o âmbito prosódico, responsável por organizar a estrutura informacional, veicular as ilocuções, entre outras funções. A classificação das disfluências pode ocorrer em dois níveis, e seus tipos nem sempre são coincidentes. O primeiro nível pode ser chamado nível locutivo, pois diz respeito aos aspectos estruturais da fala, independentemente de seu valor funcional. O segundo nível será chamado informacional-ilocucionário, que diz respeito aos aspectos mais altos do significado pragmático. Uma disfluência que se mostra como uma repetição de palavras lexicais no nível locutivo pode ser compreendida como uma mudança de planejamento informacional no nível pragmático. Consideramos bastante razoável a ideia de

que disfluências no nível informacional geram diversas consequências no nível locutivo. Tais consequências também podem ser compreendidas e classificadas como disfluências nesse nível. A possibilidade desses fenômenos ocorrerem nessas duas instâncias (estrutural e pragmática) torna insuficiente uma classificação baseada apenas no âmbito estrutural, que muitas vezes sofre consequências de reelaborações, reparos e mudanças de planejamento no âmbito pragmático.

Evidentemente, vários casos de disfluência podem ocorrer de forma restrita ao nível locutivo, sem repercussões na organização informacional ou na ilocucionária. Um reparo de itens lexicais, por exemplo, pode ser compreendido em ambos os níveis como um reparo, isto é, o falante usou uma palavra ao invés de outra por um mero acidente de seleção lexical ou por não ter tido as condições de pensar numa opção melhor antes da execução. Nesse caso, o tipo de disfluência no nível locutivo coincide com aquele do nível informacional, já que tanto a função pragmática quanto a função estrutural do reparo lexical são a mesma: substituir um ou mais itens lexicais por outro(s). Chamo a atenção, no entanto, que ambos os níveis (locutivo e informacional-ilocucionário) estão sujeitos a todos os tipos de disfluência propostos em nossa classificação. O reparo pode ocorrer para substituir uma ilocução por outra, e seu processo pode resultar em diferentes tipos de disfluência no nível locutivo, como repetição de palavras, escansões, pausa preenchida, etc.

Parecem também ocorrer dois níveis de disfluência dentro do próprio nível locutivo. Esses casos são mais evidentes em algumas unidades escansionadas. As escansões são compreendidas tradicionalmente como a quebra do isomorfismo entre uma unidade informacional e uma unidade entonacional. Quando há uma escansão, o falante veicula uma unidade informacional em duas ou mais unidades entonacionais. Isso pode ocorrer por variados motivos, inclusive com funções retóricas a serem descritas. No geral, a quebra do isomorfismo pelas escansões deve ser compreendida como uma falha de execução (Cresti, 2000), ou seja, um tipo específico de disfluência em que o falante não teve condições de terminar a unidade informacional dentro dos limites de uma única unidade entonacional; ele então acrescenta uma fronteira prosódica não terminal que escansiona a unidade pragmática em duas ou mais unidades entonacionais.

Nossa proposta considera que as escansões podem ocorrer (1) sem interferência no conteúdo locutivo da unidade - apenas com a quebra não terminal - ou, em alguns casos, (2) condicionando a ocorrência de outras disfluências no mesmo nível, como repetições e reparos. Isso fica claro quando observamos repetições de itens funcionais logo após os casos daquilo que chamamos de *scannings*.

Exemplo 64 *bfamnn01* C-ORAL-BRASIL I

*MAI: ea nũ andava no [/1]=SCA=@scanning_unclear@ no limpo não /=CMM= era só dentro do mato //CMM=

O exemplo 64 ilustra um desses casos, em que o falante, por motivo incerto, adiciona uma fronteira não terminal antes da finalização da primeira unidade informacional. A *scanning* subdivide a unidade ilocucionária em duas unidades entonacionais. A repetição da contração “no”, nesse caso, pode ser interpretada como uma consequência da *scanning* para manter, na mesma unidade entonacional, a preposição que rege o sintagma preposicionado “no limpo”. Não é estranho pensar, então, que mesmo disfluências do nível locutivo (como as *scannings*) possam condicionar disfluências desse mesmo nível (como as repetições). Algo análogo a isso foi proposto por Schegloff (2013), com seus reparos de primeira e de segunda ordens. Por enquanto, nossa anotação - assim como a de Schegloff - só fornece meios para registrarmos a disfluência geradora.

A proposta pode então ser sumarizada desta forma. Há dois níveis nos quais as disfluências podem ocorrer: o nível locutivo e o nível informacional-ilocucionário. Aquelas ocorridas no nível locutivo não têm implicação no nível informacional-ilocucionário; aquelas ocorridas no nível informacional-ilocucionário parecem sempre condicionar disfluências no nível locutivo. Nesse segundo caso, nem sempre há coincidência entre o tipo de disfluência ocorrido no nível informacional-ilocucionário com aquela que é sua consequência no nível locutivo. No âmbito no nível locutivo, as escansões podem servir como geradoras de outras disfluências, além de poderem ocorrer como uma disfluência isolada. Vejamos outro caso no exemplo 65.

Exemplo 65 *bfamnn01* C-ORAL-BRASIL I

*DFL: &he /=TMT=@filled_pause_prospection@ &el [/1]=EMP=@repair_sense_of_inadequacy@ papai foi o irmão que mais deu [/1]=SCA=@scannin_unit_prospection@ deu apoio a ele //COM=

Nesse exemplo, temos três disfluências que ocorrem em sequência. A primeira se trata de uma pausa preenchida (*filled pause*), que é causada pela tentativa de a falante preencher o turno conversacional enquanto planeja a próxima unidade informacional. A pausa preenchida é consequência do planejamento informacional e é materializada no conteúdo locutivo através de segmentos geralmente vocálicos contínuos. A segunda disfluência do exemplo se trata de um reparo lexical, que se deu quando a falante realiza parte dos segmentos da palavra “el[e]” e logo

em seguida a substitui por “papai”, na unidade entonacional seguinte. Esse reparo atua apenas no nível lexical, substituindo uma palavra por outra sem impactar o planejamento informacional ou a veiculação da ilocução. Ele parece ocorrer pelo fato de que a falante julgou inadequada a recuperação do referente apenas por meio do indexical “ele”.

A terceira disfluência do exemplo 65 se trata na verdade de duas: no curso da unidade ilocucionária, a falante produz intencionalmente uma quebra não terminal, escansionando a unidade informacional em duas unidades entonacionais. Isso ocorre provavelmente em virtude de algum tipo de dificuldade que ela experienciou na execução da unidade. Quando o fluxo é retomado após a quebra prosódica, há a repetição do verbo “deu”, que já havia sido realizado na unidade entonacional anterior. Essa repetição não parece ser a disfluência primária nesse trecho, mas uma consequência da ruptura do isomorfismo em virtude da dificuldade de planejamento. Como uma tentativa de iniciar a unidade entonacional sem implicar na separação entre o núcleo do sintagma verbal (“deu”) e seu complemento (“apoio a ele”), ela opta por repetir o verbo ao invés de fragmentar ainda mais a estrutura sintática.

Exemplo 66 *bfamnn01* 

*DFL: eu falei /=INT= eh /=CMM_r= pois é /=CMM_r= mas ele é igual [/3]=SCA_r=
@change of planning_program alteration@ pa mim /=TOP_r= ele é igual a dos outros <hhh> /=COM_r=

O exemplo 66 já foi visto na seção anterior, mas conta com um caso bastante curioso do ponto de vista dos níveis locutivo e informacional-ilocucionário. A quarta unidade entonacional, planejada como unidade ilocucionária, é apenas parcialmente realizada: “mas ele é igual...” em função de uma mudança de planejamento informacional. Ao longo da execução, a falante interrompe a unidade em construção com uma quebra prosódica não terminal; na sequência, produz uma unidade de tópico e finaliza o enunciado com a unidade ilocucionária completa - a mesma que estava sendo construída antes da interrupção para a realização do tópico. Houve dois fenômenos, portanto, em níveis diferentes: no âmbito informacional, há a mudança de planejamento motivada pela alteração no programa, já que houve a necessidade de se estabelecer de forma precisa um contexto cognitivo (através da unidade de tópico) para a aplicação da ilocução. No âmbito locutivo, há a repetição da estrutura parcialmente realizada (“ele é igual”) na unidade anterior e interrompida para a execução do tópico. Essa repetição funciona como uma consequência estrutural da mudança de planejamento informacional. Esse é um caso em que os dois tipos não coincidem. Nota-se também que a repetição não envolveu a conjunção adversativa “mas”, que não deve ser interpretada de modo composicional ao tópico

(como sugere a anotação informacional realizada), mas sim à unidade de comentário efetivamente realizada. Por esse motivo, inclusive, a anotação da unidade fragmentada como SCA é insuficiente para garantir a composicionalidade da conjunção à unidade de COM.

Essa proposta de compreensão das disfluências em dois níveis é condizente com o esquema dos padrões informacionais definidos pela L-AcT. Considerar que todas essas operações de disfluência podem impactar tanto a organização final da estrutura informacional quanto suas composições locutivas permite vislumbrar de forma ampla o potencial de ação dos fenômenos de fragmentação. Além disso, essa abordagem é, em alguma medida, compatível com a etiologia das disfluências no modelo de verbalização da experiência em Chafe. Da mesma forma que as disfluências em geral emergem das dificuldades com as operações de ajuste entre o pensamento e a linguagem, as disfluências do nível locutivo surgem pela reorganização causada por aquelas realizadas no nível informacional-ilocucionário.

5. Considerações finais

Neste trabalho, apresentamos uma classificação *corpus-driven* de disfluências baseada na estrutura informacional da fala espontânea. Partimos da compreensão etiológica desses fenômenos para propormos um conjunto fechado de tipos de disfluência e tipos de motivação. A partir da anotação de uma amostra do minicorpus do C-ORAL-BRASIL I, foi feita uma análise *corpus-based* da frequência de cada uma das categorias em face das unidades informacionais propostas pela *Language into Act Theory*. Por fim, interpretamos a ocorrência de disfluências da fala a partir de dois níveis, estrutural e pragmático.

A classificação *corpus-driven* conta com 7 categorias de disfluência (*scanning, repair, repetition, abandonment, filled pause, change of textual planning* e *change of informational planning*) e 9 categorias de motivação (*prospection, external interrupt, program alteration, inadequacy, articulation problem, improvement, immediate response, uncertainty, satisfaction*). A anotação foi realizada manualmente num conjunto de 10 textos provenientes do minicorpus. Os resultados quantitativos mostram que determinados tipos de disfluência ocorrem junto a tipos específicos de motivações. Por exemplo, as *filled pauses* se associam fortemente à *prospection*, enquanto que os *abandonments* ocorrem motivados por *external interrupt*. Os tipos com maior frequência relativa são *scanning, repair* e *repetition*, que juntos representam 64% dos casos anotados.

A análise das disfluências segundo a arquitetura do corpus revelou que os monólogos contam com maiores quantidades de *scanning, repair, repetition* e *filled pause*. Disfluências classificadas como *abandonments* ocorrem mais frequentemente em conversações, tipologia em que o turno é mais disputado pelos falantes. Exceto para as *scannings*, o contexto privado favorece de forma consistente a maior frequência de todos os tipos de disfluência. As duas variáveis da arquitetura, tipologia e contexto, condicionam de modo diferente a frequência dos fenômenos. Enquanto que a tipologia interfere na quantidade de tipos específicos de disfluência, os dois contextos influenciam a frequência global independentemente do tipo.

O cruzamento dos dados com a frequência de unidades informacionais revelou alta ocorrência relativa de disfluências em Parentéticos, Comentários Ligados e Tópicos. Essas unidades envolvem habilidades pragmáticas e/ou cognitivas destacáveis em relação às outras. Quando são longos, os Parentéticos envolvem alterações dêiticas e a suspensão temporária do nível ilocucionário. A alta frequência de disfluências em Comentário Ligado pode estar relacionada ao papel de elaboração textual desempenhado por essa unidade, já que a maior preocupação do falante nas stanzas é desenvolver o conteúdo locutivo, e não realizar ações com

alta variabilidade de classes ilocucionárias. De forma análoga, os Tópicos são unidades essenciais para o *displacement*, possibilitando que as ilocuções sejam interpretadas para além do contexto imediato da interação. Pela sua própria indexicalidade ao comentário, a realização do Tópico exige planejamento textual e boa execução do núcleo prosódico. Esses fatores pragmáticos associados a essas unidades parecem ser suficientes para condicionarem nelas maior quantidade de disfluência. Embora sejam bastante frequentes, é muito raro encontrar disfluências em unidades dialógicas ou em Comentários Múltiplos. Mesmo que os CMMs também envolvam função ilocucionária, é possível pensar que o alto nível de convencionalização entre as unidades envolvidas diminui significativamente as chances de problemas de planejamento e execução.

Além dos resultados quantitativos, propusemos uma interpretação da ocorrência dos fenômenos em face das camadas pragmática e estrutural. Parece razoável que o escopo das disfluências possa ser tanto o nível locutivo quanto o informacional-ilocucionário. Quando atuam sobre o nível locutivo, são feitas reelaborações que não impactam as funções pragmáticas da sequência linguística; quando o fazem sobre o nível informacional-ilocucionário, são realizadas reelaborações na base do que se deseja comunicar. As disfluências ocorridas nesse nível podem criar outras disfluências no nível locutivo, geradas pela necessidade de acomodação da estrutura linguística à intenção reelaborada. A quebra do isomorfismo nas *scannings* deve ser interpretada como disfluência de natureza prosódica, que, por sua vez, também pode ocasionar outras disfluências no nível locutivo, como as repetições de itens funcionais. Essa interpretação das *scannings* não deve ser confundida com aquela utilizada em *retractings* de cancelamento parcial (*retractings*-SCA). Esses últimos podem comportar uma série de fenômenos de disfluência de natureza diversa. Penso que a noção de *retracting* deve ser melhor compreendida como técnica de transcrição, e não como um tipo específico de disfluência.

Este trabalho abarcou tanto a criação de classes ontológicas a partir de dados empíricos quanto forneceu explicações etiológicas para os fenômenos classificados. A complementaridade entre a proposta da verbalização da experiência em W. Chafe e a descrição da estrutura informacional na *Language into Act Theory* foi o que possibilitou o tratamento das disfluências nos âmbitos da estrutura e da função. As categorias criadas merecem ser revisitadas à luz de novos dados, haja vista a realidade multifacetada desses fenômenos.

APÊNDICE A

Exemplos e Material Complementar

[HERE](#)

APÊNDICE B

O estatuto gramatical das pausas preenchidas na fala espontânea: uma descrição do português brasileiro e do italiano baseada em corpora

Introdução

O objetivo deste trabalho é continuar uma descrição acústica e informacional de pausas preenchidas (PP) em corpora de fala espontânea que documentam as cinco principais línguas românicas (português brasileiro, português europeu, espanhol, italiano e francês). Até o momento, os dados foram analisados para o português brasileiro e para o italiano. Investigações dessa natureza fazem parte do rol de estudos sobre disfluências, elementos caracterizados como interrupções da sequência linguística por conta de falhas em seu planejamento, erros de produção e hesitações (Kosmala, 2024).

As pausas preenchidas podem ser compreendidas como instâncias em que o falante preenche parte de seu turno conversacional com um ou mais segmentos que não contribuem com a composicionalidade morfossintática e/ou semântica do enunciado, embora sejam realizados intencionalmente. Investigações psicolinguísticas clássicas interpretam as PPs como consequências de problemas de planejamento, sem associá-las a propriedades gramaticais de sistemas linguísticos específicos (Levelt, 1984; Hartsuiker, 2009). Por outro lado, trabalhos contemporâneos sugerem a gramaticalização desses fenômenos, assim como de outros tipos de disfluências (Kosmala, 2024; Dingemanse, 2024). O estudo de Kirjavainen (2022) observou para o inglês que PPs produzidas apenas com segmentos orais (“uh”) ocorrem em contextos morfossintáticos diferentes daquelas produzidas com segmentos nasais (“uhn”) no BNC corpus.

Os exemplos 1 e 2 abaixo apresentam pausas preenchidas em português brasileiro e em italiano para sequências linguísticas percebidas como prosodicamente terminadas. No sistema CHAT de anotação, usado pelos corpora da família C-ORAL, as PPs são transcritas como **&he**, independentemente do segmento realizado. As barras simples “/” representam limites de unidades entonacionais, enquanto que as barras duplas “//” representam unidades entonacionais em final de sequência terminada (Raso & Mello, 2012).

Exemplo 1 C-ORAL-BRASIL I (bfamnn01) [\[ÁUDIO\]](#)

*GIL: nós estamos nesse momento / **&he** / &fa [1] **&he** / procurando [1] **&he** / **&he** / no [1] no processo / de procurar / **&he** / **&he** / **&he** / quadra / olhar os melhores lugares //

Exemplo 2 C-ORAL-ROM (ifammn16) [\[ÁUDIO\]](#)

*ROB: le piaghe da decubito / **&he** / sono da tenere sotto controllo //

Em 1, são realizadas 7 (sete) PPs com vogais média-baixa anterior ([ε]), diferentes durações e contextos morfossintáticos. No exemplo 2, ocorre uma PP realizada em uma combinação entre consoante e vogal nasal. Neste trabalho, estamos interessados em descrever a estruturação acústica de pausas desse tipo para diferentes línguas românicas através de um estudo de corpora (C-ORAL-BRASIL e C-ORAL-ROM).

O texto está organizado da seguinte forma: a próxima seção trata da relação entre as unidades entonacionais, os focos de pensamento (Chafe, 1994; Cresti, 2000) e algumas consequências etiológicas para a existência de disfluências. Na sequência, serão descritos os corpora e a metodologia usada para o estudo das pausas preenchidas. Por fim, estarão dispostos os resultados e as discussões para os dados do português brasileiro e do italiano.

Disfluências e o foco do pensamento

A fala espontânea pode ser segmentada de variadas maneiras, que dependem de interesses teóricos e condições práticas. Tradicionalmente, a segmentação pode se dar através de limites sentenciais (como a máxima projeção do sintagma verbal), por momentos de pausas silenciosas (quando se tem acesso aos arquivos de áudio) e de muitas outras formas (Izre'el et al., 2020). A questão que coloca em xeque a adequação de cada uma dessas segmentações diz respeito ao pareamento função-entonação, que muitas vezes não condiz com momentos de silêncios e nem respeita fronteiras sintáticas.

W. Chafe (1994) desenvolve a ideia de que as chamadas unidades entonacionais representam a unidade básica da fala. Ao falar, o indivíduo seleciona focos de pensamento que são verbalizados através das unidades entonacionais. Estas, por sua vez, possuem contornos prosódicos bem marcados (clara excursão de f0, duração e intensidade características) que permitem a identificabilidade das fronteiras prosódicas através da percepção. A existência de unidades entonacionais e a capacidade perceptual dos falantes podem ser comprovadas por diferentes experimentos estatísticos. O mais comum deles é o teste *inter-rater agreement*, no qual um grupo de pessoas diferentes entrará em contato com um corpus e fará segmentações da fala espontânea. Em geral, há mais de 80% de similaridade entre as segmentações, mostrando

que os falantes possuem alta concordância na identificação de padrões prosódicos (Panunzi; Gregori; Rocha, 2020).

Enquanto Chafe associa as unidades entonacionais aos focos de consciência, Cresti (2000) propõe uma associação baseada na estrutura informacional na *Language into Act Theory (L-Act)*. Partindo na *Information Patterning Hypothesis* (t'Hart et al., 1990), Cresti sugere que as unidades entonacionais (portanto, com forma prosódica delimitada) estão relacionadas biunivocamente a unidades informacionais (portanto, com funções delimitadas). Nesse ponto de vista, as unidades possuem formas prosódicas específicas que se convencionalizaram para desempenhar determinadas funções informacionais. As relações morfossintáticas são realizadas majoritariamente dentro dos limites de unidades informacionais e só podem ser interpretadas pelo ouvinte após o reconhecimento da estrutura informacional (Cresti, 2000; Cresti & Moneglia, 2005; Bossaglia, 2015).

A figura 1 representa um enunciado realizado em duas unidades entonacionais; cada uma delas possui uma forma prosódica e corresponde a uma unidade informacional específica. A estrutura morfossintática do enunciado é simples, mas a dinâmica prosódico-pragmática é surpreendentemente complexa. A primeira unidade conta com uma forma prosódica de Tópico, que, no âmbito da L-AcT, é uma unidade responsável por delimitar um âmbito cognitivo dentro do qual uma ilocução será aplicada. A unidade seguinte, “tão aqui”, constitui a própria ilocução¹⁷. Herdando da Teoria dos Atos de Fala (Austin, 1962) o caráter acional da linguagem, Cresti propõe que a fala espontânea se constitui do mais alto grau de troca de ações. A ilocução representa uma ação individual baseada em uma relação de *aboutness* pragmático entre o enunciado e o interlocutor (Rocha, em preparação).

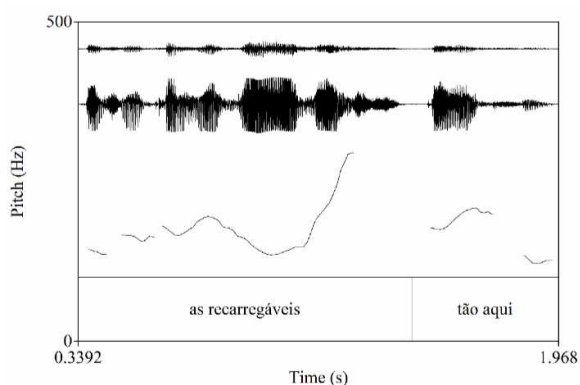


Figura 1. Excursão de f0 de um enunciado do C-ORAL-BRASIL I. Fonte: Rocha & Vital (2025).

¹⁷ As unidades informacionais que realizam ilocuções também possuem formas prosódicas. Uma proposta para a diferenciação entre ilocuções e atitudes com base na forma pode ser apreciada em Raso & Rocha (2021).

O tema das disfluências tem sido abordado por diferentes linhas teóricas e produz argumentos importantes para a tese de que a linguagem não atua como estruturadora do pensamento. Na verdade, Chafe (1994) coloca as disfluências da fala espontânea como uma evidência de que os falantes possuem consciência de mais do que aquilo que é dito ao longo do discurso; é precisamente pelo fato de o pensamento se dar independentemente da fala que o falante tem a possibilidade de se autocorriger e de fazer reelaborações. É esse movimento causado pelo ajuste pensamento-linguagem que faz com que a produção da fala seja enquadrada como um ato de criação:

Hesitation phenomena [...] provide good evidence that speaking is not a matter of regurgitating material already stored in the mind in linguistic form, but that it is a creative art, relating two media, thought and language, which are not isomorphic but require adjustments and readjustments to each other. A speaker does not follow a clear, well traveled path, but must find his way through territory not traversed before, where pauses, changes of direction, and retracing of steps are quite to be expected. The fundamental reason for hesitating is that speech production is an act of creation. (Chafe, 1980, p. 170.)

Pelo seu processamento online, a fala exige momentos de pausas, reelaborações e autocorrekções. As disfluências (ou fenômenos de hesitação, na nomenclatura da *Thought-based linguistics*) podem ser entendidas como consequências dos processos de seleção, categorização, orientação e combinação para a verbalização da experiência (Chafe, 1980; Chafe, 2005).

A proposta de Chafe ecoa na L-AcT, embora o aspecto das disfluências tenha sido melhor elaborado nos últimos anos. A partir de estudos *corpus-driven*, Vital; Rocha; Raso (2025) propõem uma classificação desses fenômenos ancorada na estrutura informacional, cujas funções são acessíveis e estão empiricamente descritas. Enquanto Chafe fundamenta a natureza etiológica das disfluências, os trabalhos com a L-AcT se propõem a uma classificação baseada no uso, procurando por padrões formais que podem indicar a gramaticalização desses elementos no âmbito de diferentes línguas. Até o momento, foram observadas seis classes que se diferenciam com relação à continuidade ou interrupção da unidade informacional (IU) em curso no momento da disfluência (tabela 1). Nesse panorama, um reparo (*repair*), por exemplo, passa a ser descrito como quando o falante faz autocorrekções dentro da mesma unidade informacional. Essas correções podem envolver variados níveis linguísticos, como reparos lexicais (ex.: na alteração de itens lexicais), fonéticos (ex.: reparo de pronúncia), morfossintáticos (ex.: reparo de concordância), semântico (ex.: reparo de tempo verbal) e pragmático (ex.: reparo de ilocução).

Tabela 1. Classes de disfluências identificadas por Vital; Rocha; Raso (2025) através de estudo *corpus-driven*.

Continuity of the IU	Class
Continuity	Repair
	Repetition
	Change of planning
	Scanning unit
	Filled pause
No Continuity of the IU	Abandonment
	Change of planning

Nessa caracterização, as pausas preenchidas (*filled pauses*) constituem uma classe de disfluências que ocorrem dentro de uma unidade informacional. Na possibilidade de aliarmos a interpretação estrutural da L-AcT à abordagem etiológica de Chafe, é possível interpretar as PPs como consequências dos processos de ajuste pensamento-linguagem que ocorrem durante a estruturação de uma unidade informacional. Enquanto o falante precisa de um breve período de tempo para planejar a estrutura linguística da unidade, ele a preenche com diferentes segmentos (as pausas preenchidas) que, embora não contribuam com a proposição, se convencionalizaram para expressarem funções de manutenção do turno conversacional.

Na próxima seção, descreveremos os métodos usados para a descrição acústica das PPs, levando em consideração os pressupostos teóricos expostos acima: as disfluências nascem do ajuste pensamento-linguagem (Chafe) e se diferenciam com relação à continuidade ou não da unidade informacional (Cresti).

Métodos

A metodologia empregada neste trabalho é amplamente baseada nas observações de Kirjavainen et al. (2022) acerca de PPs na língua inglesa. Os autores analisaram integralmente o BNC corpus em busca de diferenças estruturais entre pausas com qualidades nasais e pausas com qualidades orais. No entanto, ao invés de usarem as gravações de áudio disponíveis, o trabalho foi pautado na busca automática de “uh” (pausa oral) e de “uhn” (pausa nasal) dentre os arquivos de transcrição. Foram feitas análises de colocados e de contextos morfossintáticos após anotação PoS. Acreditamos, no entanto, que a revisão atenta de cada uma das ocorrências de pausas não deva ser desconsiderada num trabalho dessa magnitude. Mesmo em um corpus de fala com alta consistência e acurácia entre os transcritores, a revisão dos dados por revisores experientes é essencial para manter a consistência dos resultados finais.

Os corpora utilizados neste trabalho são o C-ORAL-ROM (Cresti & Moneglia, 2005) e o C-ORAL-BRASIL (Raso & Mello, 2012). Ambos são corpora de referência para fala espontânea e contam com segmentação prosódica e alta variação diafásica. O C-ORAL-ROM compila gravações das principais línguas românicas - português europeu, espanhol, francês e italiano - e conta uma interface operável online. O C-ORAL-BRASIL I é o membro brasileiro da família dos corpora C-ORAL e conta com 139 gravações espontâneas em português brasileiro transcritas e segmentadas prosodicamente. Junto dos corpora, são disponibilizados os arquivos de áudio em formato .wav, além dos arquivos de alinhamento texto-som, metadados das gravações e transcrições semiortográficas.

A arquitetura dos corpora é pautada em diferentes contextos (familiar e público) e tipologias (monológica, dialógica e conversacional), e o tamanho médio das gravações é de cerca de 1.500 palavras. Para o estudo das pausas preenchidas, foram selecionadas, de forma a balancear os contextos e as tipologias, gravações do C-ORAL-BRASIL I (N = 40 gravações) e da parte italiana do C-ORAL-ROM (N = 11 gravações), conforme a tabela 2. Nas próximas etapas, o mesmo número selecionado para o italiano será selecionado para as outras línguas do C-ORAL-ROM, de modo a se conseguir uma quantidade comparável estatisticamente àquela presente no C-ORAL-BRASIL.

Tabela 2. Número de gravações selecionadas por amostragem aleatória para análise de PPs em PB e IT.

PB	familiar	público	IT	familiar	público
conversação	8	4	conversação	2	1
diálogo	10	4	diálogo	3	1
monólogo	12	2	monólogo	3	1

Uma vez selecionadas as gravações, os trechos em texto e em áudio com pausas preenchidas, transcritas como &he, foram isolados automaticamente através de script em Python. Todos os trechos foram analisados manualmente por dois revisores experientes. A revisão consistiu na checagem da transcrição, atestação ou não da PP e anotação das seguintes propriedades acústicas:

1. Segmento imediatamente anterior (quando presente)
2. Segmento principal
3. Segmento imediatamente posterior (quando presente)
4. Presença de *creaky voice*
5. Presença de traço nasal

6. A estrutura silábica das pausas preenchidas

Para o português brasileiro, foram contabilizadas 199 PPs nos 40 textos; para o IT, foram contabilizadas 126 PPs nos 11 textos. Silêncios não preenchidos e alongamentos de sílabas não foram considerados. Para 20 textos do C-ORAL-BRASIL, foram contabilizadas as unidades informacionais mais frequentes à esquerda (encerradas antes da PP) e à direita (unidade dentro da qual a PP é realizada). A próxima seção descreve os resultados para o PB e para o IT.

Resultados

Os resultados serão apresentados comparativamente seguindo o fluxo de propriedades acústicas anotadas e descritas na metodologia. Para a análise das vogais que constituem o segmento principal nas PPs, o PB contou com 5 grupos de segmentos (ϵ , $\text{\textcircled{a}}$, e , nasais, outros) (figura 2), versus 11 grupos em IT (ϵ , em , m , $\text{\textcircled{a}}$, e , em , $\text{\textcircled{a}}$, $\text{\textcircled{a}}$, am , \tilde{a} , a) (figura 3). Para ambas as línguas, a vogal [ϵ] é a mais frequente em pausas preenchidas, representando 76,9% das pausas em PB e 25,4% em IT.

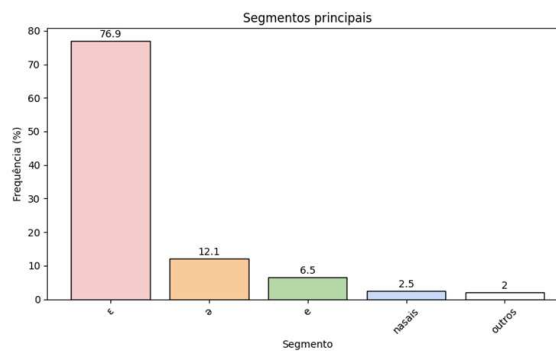


Figura 2. Vogais mais frequentes em pausas preenchidas do PB.

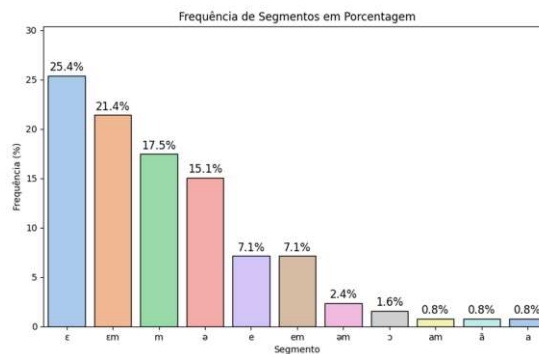


Figura 3. Vogais mais frequentes em pausas preenchidas do IT.

A análise de segmentos nasais e segmentos orais também demonstrou diferença importante entre as duas línguas. Embora o traço de nasalidade tenha representatividade fonológica no português, ele ocorreu com pouca frequência nos dados de PPs (6,0%) (figura 4). No italiano, por outro lado, a ocorrência de segmentos nasais não fornece distinção fonológica na língua, mas representa a maior parte das observações em pausas preenchidas (50,8%) (figura 5). O mesmo foi observado para a qualidade *creaky*, que é mais frequente no italiano do que no português.

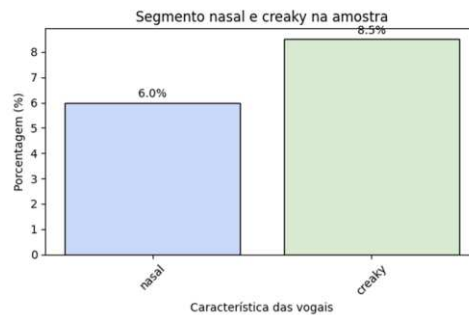


Figura 4. Frequência de segmentos nasais e creaky voice em PPs no PB.

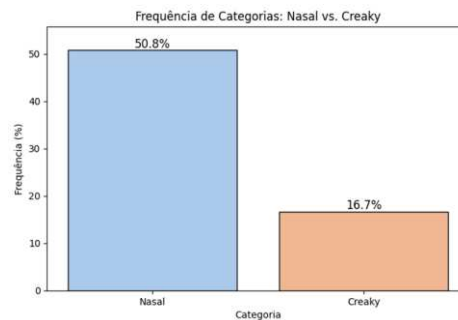


Figura 5. Frequência de segmentos nasais e creaky voice em PPs no IT.

Para ambas as línguas, foi contabilizada também a frequência de PPs com relação ao contexto e a tipologia. Não houve diferença na distribuição; por conta disso, os dados foram juntados num único gráfico (figura 6). Os cruzamentos *familiar-monológico* e *familiar-público* foram aqueles que contaram com a maior quantidade de PPs. Diferentemente das outras tipologias, o monólogo comporta uma estrutura conversacional em que o falante tem menos concorrência de turno de fala. É isso o que permite a maior elaboração estrutural e informacional (Raso & Mello, 2012) presente nessas interações. Com a necessidade constante de ajuste de pensamentos complexos à estrutura limitada da linguagem no processo de verbalização, os falantes produzem muitas PPs.

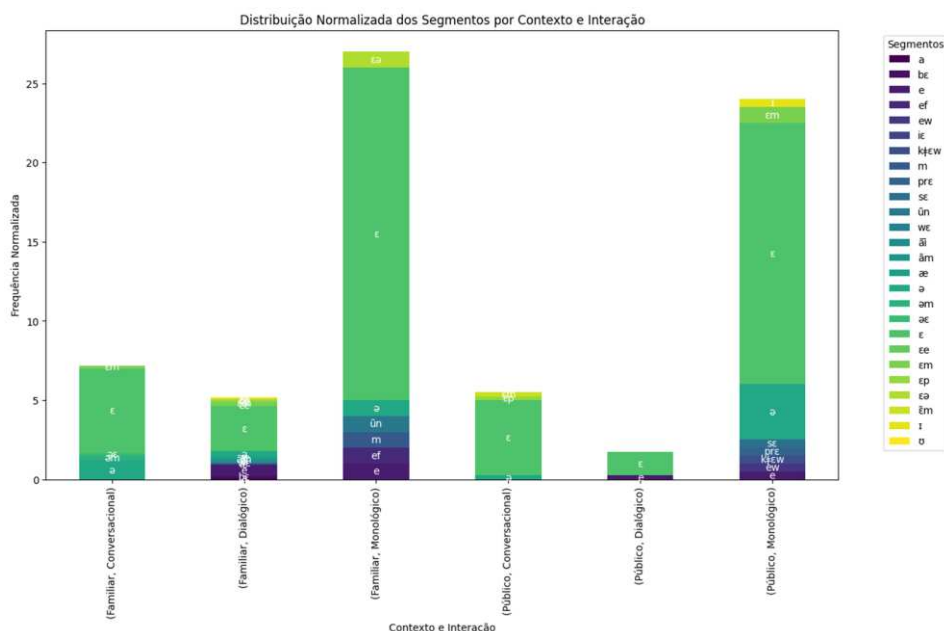


Figura 6. Frequência de PPs em PB e IT no cruzamento de contextos (familiar, privado) e tipologias (monológica, dialógica, conversacional).

Quanto à estruturação silábica das pausas preenchidas, em ambos os corpora a formação V (pronúncia de apenas uma vogal) foi mais frequente, representando 65,1% no C-ORAL-BRASIL I e 53,9% no C-ORAL-ROM. Em segundo lugar, as pausas constituídas por estrutura VC, sendo que a consoante nasal representa 3,3% das ocorrências nesse contexto para o PB e 56,2% no IT.

Discussão

Neste trabalho, inicialmente cotejamos a interpretação das disfluências partindo de duas propostas complementares: a de Chafe (1980; 1994; 2005) e a de Cresti (2000) através da L-AcT. Chafe nos fornece uma proposta etiológica para a natureza dos fenômenos de disfluências, enquanto a L-AcT possibilita uma classificação ontológica desses elementos a partir da continuidade ou não das unidades informacionais. Para o cumprimento dos objetivos, fizemos um recorte de um dos tipos de disfluências descritas pelo trabalho de Vital; Rocha; Raso (2025): as pausas preenchidas. Foram analisadas a composição acústica e pragmática de 199 PPs para o PB e 126 para o IT.

Os resultados frequenciais das PPs complementam os achados de Kirjavainen et al. (2022), já que a forma convencional desses elementos em português e em italiano é significativamente diferente. Em italiano, há um conjunto maior de segmentos utilizados nas PPs do que aqueles disponíveis no português. A estrutura silábica mais frequente, uma única

sílaba preenchida por uma vogal (V), é a mesma para as duas línguas. Embora o português conte com distinção fonêmica entre vogais orais e nasais, o uso do traço nasal é escasso na produção de pausas. Isso não ocorre em italiano, em que mais da metade das ocorrências conta com traço nasal. Resultados como esse apontam para uma gramaticalização desses itens - antes considerados idiossincráticos e aleatórios - sob pressão do sistema fonológico das línguas.

Este trabalho continuará em desenvolvimento, e o comportamento das PPs nas outras línguas românicas disponíveis no C-ORAL-ROM será descrito no alinhamento da abordagem de Chafe e de Cresti.

6. Referências

- ALBUQUERQUE, M. *Fala espontânea e gesto: um estudo qualitativo baseado em corpus*. Dissertação de Mestrado - Universidade Federal de Minas Gerais, 2025.
- ALLWOOD, J.; NIVRE, J.; AHLSEN, E. Speech Management - On the Non-written Life of Speech. *Nordic Journal of Linguistics*, 13(1), 3–48. 1990.
- AUSTIN, J. L. *How to do things with words*. Oxford University Press, Oxford, 1962.
- BETZ, S. *Hesitations in Spoken Dialogue Systems*. Tese de Doutorado - Bielefeld University. 2020.
- BICK, E. A anotação gramatical do C-ORAL-BRASIL. In: RASO, T.; MELLO, H. *C-ORAL-BRASIL I: Corpus de Referência do Português Brasileiro Informal*. Belo Horizonte: Editora UFMG, p. 223-254, 2012.
- BLANCHE-BENVENISTE, C. *La langue parlée. Le grand livre de la langue française*, p. 317-344, 2003.
- BOSSAGLIA, G. Orientação pragmática da sintaxe na fala espontânea: uma análise corpus-based da subordinação completiva e adverbial no português do Brasil. *Domínios de Linguagem*, v. 9, n. 5, p. 309-335, 2015. <https://doi.org/10.14393/DLE-v9n5a2015-16>
- BOSSAGLIA, G.; CAVALCANTE, F., & RASO, T. Minicorpus de português brasileiro falado informal. 2018.
- CÂMARA-JR, J. *Estrutura da língua portuguesa*. Rio de Janeiro: Vozes, 1970.
- CAVALCANTE, F. A., *The topic unit in spontaneous american English: a corpus-based study*. Master's dissertation in Linguistics - Universidade Federal de Minas Gerais, 2016.
- CAVALCANTE, F. A., *The informational unit of topic: a crosslinguistic, statistical study based on spontaneous speech corpora*. PhD dissertation in Theoretical and Descriptive Linguistics, Universidade Federal de Minas Gerais, 2020.
- CHAFE, W. Some reasons for hesitating. In: DECHERT, W.; RAUPACH, M. (Ed.). *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler*. Walter de Gruyter, p. 169-180, 1980.
- CHAFE, W. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In: LI, C. *Subject and topic*, New York: Academic Press, 1976.
- CHAFE, W. *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. University of Chicago Press, 1994.
- CHAFE, W. Putting grammaticalization in its place. In: WISCHER, I.; DIEWALD, G. *New reflections on grammaticalization*. John Benjamins Publishing Company, 2002. p. 395-412.

CHAFE, W. The relation of grammar to thought. In: BUTLER, C.; GÓMEZ-GONZÁLEZ, M.; DOVAL-SUÁREZ, S. (Ed.). *The Dynamics of language use: Functional and contrastive perspectives*. John Benjamins Publishing Company, 2005. p. 57-78.

CHAFE, W. *Meaning and the Structure of Language*. Chicago University Press, 1970.

CLARK, H. *Using language*. Cambridge University Press, 1996.

CRESTI, E. *Corpus di Italiano Parlato. Introduzione e Campioni*. Accademia della Crusca, 2000.

CRESTI, E. La Stanza: un'unità di costruzione testuale del parlato. In: *Atti del X Congresso della Società Internazionale di Linguistica e Filologia Italiana*, SILFI, 2008. Basilea, 2009, p. 1-25.

CRESTI, E.; MONEGLIA, M. (Ed.). *C-ORAL-ROM: integrated reference corpora for spoken romance languages*. John Benjamins Publishing, 2005.

CRESTI, E. The affective foundation of speech: affect, prosody, illocution. *Journal of Speech Sciences*. 2025. 14(00): e025005.
Available at: <https://econtents.sbu.unicamp.br/inpec/index.php/joss/article/view/20385>

CRIBLE, L.; DUMONT, A.; GROSMAN, I.; NOTARRIGO, I. (Dis)fluency across spoken and signed languages: Application of an interoperable annotation scheme. In DEGAND, L.; GILQUIN, G.; SIMON, A. (Eds.), *Fluency and Disfluency across Languages and Language Varieties (Corpora and Language in Use-Proceedings 4)*. Presses universitaires de Louvain. 2019.

DINGEMANSE, M.; ENFIELD, N. Interactive repair and the foundations of language. *Trends in Cognitive Sciences*, v. 28, n. 1, p. 30-42, 2024. <https://doi.org/10.1016/j.tics.2023.09.003>

DUBOIS, J.; CHAFE, W.; MEYER, C.; THOMPSON, S.; ENGLEBRETSON, R.; MARTEY, N. Santa Barbara corpus of spoken American English, Parts 1-4. Philadelphia: Linguistic Data Consortium, 2000-2005.

DUEZ, D. *La pause dans la parole de l'homme politique*. Editions du Centre national de la recherche scientifique, 1991.

EKLUND, R. *Disfluency in Swedish human-human and human-machine travel booking dialogues*. Tese de Doutorado. Linköping University Electronic Press. 2004

FAGIOLI, M. *Istinto di morte e conoscenza: pensieri di psicoanalisi*. Roma: Armando; 1972.

FERREIRA, F.; BAILEY, K. Disfluencies and human language comprehension. *Trends in Cognitive Sciences*, 8(5), 231-237. 2004. Disponível em: <https://doi.org/10.1016/j.tics.2004.03.011>

- GILQUIN, G. Hesitation markers among EFL learners: Pragmatic deficiency or difference. In ROMERO-TRILLO, J. (Ed.), *Pragmatics and Corpus Linguistics: A Mutualistic Entente* (pp. 119–149). De Gruyter Mouton. 2008.
- GINZBURG, J.; POESIO, M. Grammar is a system that characterizes talk in interaction. *Frontiers in Psychology*, 7, 2016. Disponível em: <https://doi.org/10.3389/fpsyg.2016.01938>
- GRICE, H. Logic and conversation. *Syntax and semantics*, v. 3, p. 43-58, 1975.
- GÖTZ, S. *Fluency in native and non-native English speech*. John Benjamins Publishing. 2013.
- GRAZIANO, M.; GULLBERG, M. When speech stops, gesture stops: Evidence from developmental and crosslinguistic comparisons. *Frontiers in psychology*, v. 9, p. 879, 2018. <https://doi.org/10.3389/fpsyg.2018.00879>
- HALLIDAY, M. Notes on transitivity and theme in English: Part 2. *Journal of linguistics*, v. 3, n. 2, p. 199-244, 1967.
- 't HART, J.; COLLIER, R.; COHEN, A. *A Perceptual Study of Intonation: An Experimental-Phonetic Approach to Speech Melody*. Cambridge University Press. 1990.
- HIEKE, A. A Content-Processing View of Hesitation Phenomena. *Language and Speech*, 24(2), 147–160. 1981.
- HOCKETT, C. *A course in modern linguistics*. Macmillan, 1960.
- KIRJAVAINEN, M.; CRIBLE, L.; BEECHING, K. Can filled pauses be represented as linguistic items? Investigating the effect of exposure on the perception and production of um. *Language and Speech*, v. 65, n. 2, p. 263-289, 2022. <https://doi.org/10.1177/13670069231216293>
- KOSMALA, L. *Beyond Disfluency: The interplay of speech, gesture, and interaction*. John Benjamins Publishing Company, 2024.
- KOWAL, S.; O' CONNELL, D. C. Pausological research at Saint Louis University. In: DECHERT, H.; RAUPACH, M. *Temporal Variables in Speech: Studies in Honour of Frieda Goldman-Eisler*, Berlin, New York: De Gruyter Mouton, 1980. Disponível em: <https://doi.org/10.1515/9783110816570>
- LEVSHINA, N. *Communicative efficiency*. Cambridge University Press, 2022.
- LEVELT, W. Monitoring and self-repair in speech. *Cognition*, v. 14, n. 1, p. 41-104, 1983.
- LICKLEY, Robin J. Fluency and Disfluency. In: REDFORD, M. *The handbook of speech production*. John Wiley & Sons, 2015. cap. 20, p. 445–474. Disponível em: <https://doi.org/10.1002/9781118584156.ch20>

- LIMA E SILVA, L. O estatuto da sintaxe na fala: considerações acerca da proposta da Language Into Act Theory. *Revista de Estudos da Linguagem*, v. 28, n. 1, 2020. <https://doi.org/10.17851/2237-2083.28.1.271-330>
- MACLAY, H.; OSGOOD, H. Hesitation phenomena in spontaneous English speech. *Word*, 15(1), 19–44. 1959.
- MACWHINNEY, B. *The CHILDES Project: tools for analyzing talk*. 3rd Edition. Mahwah: Lawrence Erlbaum Associates, 2000.
- MARTIN, P. WinPitch LTL II, a multimodal pronunciation software. In: *InSTIL/ICALL 2004 Symposium on Computer Assisted Learning*, Venice, Italy. Retrieved from http://www.isca-speech.org/archive/icall2004/iic4_042.html
- MCCARTHY, M. (2009). Rethinking spoken fluency. *ELIA*, 9, 11–29.
- MENDES, M. *As ilocuições padronizadas segundo a Teoria da Língua em Ato*. Trabalho de Conclusão de Curso, Universidade Federal de Minas Gerais, 2024.
- MELLO, H., RASO, T., MITTMANN, M., VALE, H., CÔRTEZ, P. Transcrição e segmentação prosódica do corpus C-ORAL-BRASIL: critérios de implementação e validação. In: RASO, T.; MELLO, H. *C-ORAL-BRASIL I: Corpus de referência do português brasileiro falado informal*. Belo Horizonte: Editora UFMG, 125-176. 2012.
- MONEGLIA, M.; RASO, T., Notes on Language into Act Theory (L– AcT), In: RASO, T.; Mello, H. (eds), *Spoken Corpora and Linguistic Studies*, John Benjamins Publishing Company, Amsterdam – Philadelphia 2014, pp. 468–495.
- MOREL, M.; DANON-BOILEAU, L. *Grammaire de l'intonation l'exemple du français*. Editions OPHRYS, 1998.
- PANUNZI, A.; GREGORI, L.; ROCHA, B. Comparing annotations for the prosodic segmentation of spontaneous speech: Focus on reference units. In: PANUNZI, A.; RASO, T.; MELLO, H.; IZRE'EL, S. *In Search of Basic Units of Spoken Language: A corpus-driven approach*. John Benjamins Publishing Company, 2020. p. 403-432.
- RASO, T. Prosódia e Pragmática. In: Speech Sciences Entries. Speech Prosody Studies Group, 2024. Disponível em: <https://gepf.falar.org/entries/54>
- RASO, T.; MELLO, H. (eds). *C-ORAL-BRASIL I: Corpus de referência do Português Brasileiro falado informal*, Editora UFMG, Belo Horizonte, 2012.
- RASO, T.; MELLO, H.; FERRARI, L. A.; ROCHA, B. C-ORAL-BRASIL II: corpus de referência do português brasileiro falado formal, mídia e telefone. Disponível em: http://c-oral-brasil.org/c-oral-brasil-ii_N.php
- RASO, T. Prosodic constraints for discourse markers. In: RASO, T.; MELLO, H. (Eds.). *Spoken corpora and linguistic studies*. John Benjamins Publishing Company, 2014. p. 411-467.

RASO, T.; ROCHA, B. Illocution and attitude: on the complex interaction between prosody and pragmatic parameters. *Journal of Speech Sciences*, 2017. Disponível em: <https://econtents.sbu.unicamp.br/inpec/index.php/joss/article/view/15062>

RASO, T.; ROCHA, B.; SALGADO, J.; CRUZ, B.; MANTOVANI, L.; MELLO, H. The C-ORAL-ESQ project: a corpus for the study of spontaneous speech of individuals with schizophrenia. *Language Resources & Evaluation* 58, 903–923 (2024). <https://doi.org/10.1007/s10579-023-09675-y>

ROCHA, B. The parenthetical discursive level. *Journal of Speech Sciences*. 14(00):e025007. 2025.

Disponível em: <https://econtents.sbu.unicamp.br/inpec/index.php/joss/article/view/20411>

ROCHA, Bruno. Metodologia empírica para o estudo de ilocuições do português brasileiro. *Domínios de Linguagem*, v. 7, n. 2, p. 109-148, 2013.

ROCHA, B.; RASO, T.; VITAL, A.; FERRARI, L. C-ORAL-ESQ: Corpus brasileiro de pessoas com esquizofrenia. Disponível em: <http://c-oral-brasil.org/c-oral-esq.php>

ROCHA, J. *Testing statistical methods for sociolinguistic profiling of Brazilian Portuguese speakers*. Dissertação de Mestrado. Universidade Federal de Minas Gerais. 2024.

ROCHA, J.; VITAL, Á. Comparisons between face-to-face and telephone speech: research possibilities. *Journal of Speech Sciences*. 2025; 14(00):e025003. Disponível em: <https://econtents.sbu.unicamp.br/inpec/index.php/joss/article/view/20389>

SANTOS, S. *How to deal with Discourse Markers: a prosodic, corpus-based, computational and experimental proposal*. Tese de Doutorado. Université Paris-Saclay; Universidade Federal de Minas Gerais, 2024.

SANTOS, S.; BOSSAGLIA, G. Morphosyntactic, prosodic, functional and distributional description of the information unit of Parenthesis in spoken Brazilian Portuguese. *Caletroscópio*, 2018. Available at: <https://doi.org/10.58967/caletroscopio.v6.nesp.2018.3829>

SEARLE, J. A Taxonomy of Illocutionary Acts. In: GUNDERSON, K., *Language, Mind and Knowledge*. University of Minnesota Press. pp. 344-369, 1975.

SCHEGLOFF, E. Ten operations in self-initiated, same-turn repair. In: HAYASHI, M.; RAYMOND, G.; SIDNELL, J. (Eds.). *Conversational Repair and Human Understanding. Studies in Interactional Sociolinguistics*. Cambridge University Press; 2013:41-70.

SHRIBERG, E. *Preliminaries to a theory of speech disfluencies*. Doctoral dissertation, University of California at Berkeley, 1994.

TROUVAIN, J.; CRIBLE, L.; BELZ, M.; BETZ, S.; BEŇUŠ, Š.; BAQUÉ, L.; CANTARUTTI, M.; NAPOLI, J. D.; DIDIRKOVÁ, I.; MACHUCA, M.; MAREKOVÁ, L.; NICULESCU, O.; PELTONEN, P.; PISTONO, A.; SCHETTINO, L.; SILBER-VAROD, V.; WILLIAMS, S. On variability in the identification and labelling of disfluencies - preliminary results from 23

annotations of the same data. Proc. 12th edition of the Disfluency in Spontaneous Speech Workshop (DiSS 2025), 57-61, 2025. doi: 10.21437/DiSS.2025-12

VITAL, Á.; ROCHA, B.; RASO, T. Disfluencies and Informational Structure in Spontaneous Speech: Classification and a Case Study of Filled Pauses. Proc. 12th edition of the Disfluency in Spontaneous Speech Workshop (DiSS 2025), 42-46, 2025. doi: 10.21437/DiSS.2025-9

WITTENBURG, P.; BRUGMAN, H.; RUSSEL, A.; KLASSMANN, A.; SLOETJES, H. ELAN: A professional framework for multimodality research. In *5th international conference on language resources and evaluation (LREC 2006)* (pp. 1556-1559). 2006.