

3º CICLO DE ESTUDOS EM CIÊNCIAS DA LINGUAGEM

LINGUÍSTICA

As relações causais em notícias em Português Europeu: uma análise semântica

Nana Yu

D

2026



Nana Yu

As relações causais em notícias em Português Europeu: uma análise semântica

Tese realizada no âmbito do Doutoramento em Ciências da Linguagem, orientada pela Professora Doutora Maria da Purificação Moura Silvano

Faculdade de Letras da Universidade do Porto

2026



Este trabalho está licenciado ao abrigo de uma licença CC BY 4.0

Aos meus pais, Yuehu Yu e Ronghua Cao

Ao meu irmão, Hao Yu

Sumário

Declaração de honra	7
Agradecimentos.....	8
Resumo	10
Abstract	12
Índice de Figuras	1
Índice de Tabelas	2
Lista de abreviaturas e siglas	3
Introdução	4
1.Causalidade	9
1.1. O conceito de causalidade	9
1.1.1. Abordagem filosófica da causalidade.....	12
1.1.2. Abordagem linguística da causalidade	15
1.2. Tipos de causas	26
1.2.1. Causa singular e causa explicativa (Davidson, 1967).....	27
1.2.2. Causação permissiva e causação factiva (Nedyalkov & Silnitsky, 1973).....	28
1.2.3. Causa de conteúdo e epistêmica (Sweetser, 1990; Noordman & Blijzer, 2000)	29
1.2.4. Causa direta e causa indireta (Galán Rodríguez, 1999).....	32
1.3. Conclusão.....	36
2.A Causalidade nas teorias das relações discursivas.....	37
2.1. <i>Theory of Discourse Coherence</i> (Hobbs, 1985).....	38
2.2. <i>Rhetorical Structure Theory</i> (Mann & Thompson, 1987)	43
2.3. <i>Segmented Discourse Representation Theory</i> (Asher & Lascarides, 2003)	55
2.4. Extensão da <i>SDRT</i> (Cunha & Silvano, 2009; Silvano, 2010).....	58
2.5. <i>Penn Discourse Treebank</i> (Prasad et al., 2008; 2010; 2019)	63
2.6. <i>Core-Discourse Relations</i> (ISO 24617-8, 2016).....	69
2.7. Conclusão.....	76
3.A inferência de causalidade no discurso: marcas linguísticas	78
3.1. Léxico	79
3.1.1. Verbos causativos.....	82
3.1.2. Nomes causativos.....	85
3.1.3. Advérbios com valor causal	88
3.2. Preposições	91
3.3. Conectores.....	94

3.4. Orações	101
3.5. Conclusão	110
4.As categorias de Tempo e Aspeto	112
4.1. Noção de Tempo e Aspeto	113
4.2. Algumas propostas teóricas de Tempo	115
4.2.1. Reichenbach (1947)	115
4.2.2. Kamp & Reyle (1993)	118
4.2.3. Oliveira (2003; 2013)	120
4.3. Algumas propostas teóricas de Aspeto	123
4.3.1. Vendler (1967)	124
4.3.2. Moens & Steedman (1988)	126
4.3.3. Cunha (2004; 2013)	131
4.3.4. Conclusão	134
5.Esquemas de anotação semântica de informação temporal e causal	135
5.1. Esquemas de anotação causal	136
5.2. Esquemas de anotação temporal	141
5.3. Esquemas de anotação causal e temporal	145
5.4. Conclusão	148
6.Estudo empírico das relações causais em narrativas jornalísticas	150
6.1. Descrição do estudo	150
6.1.1 Objetivos e questões de investigação	151
6.1.2 Procedimentos metodológicos	151
6.1.3 Constituição e descrição do corpus	152
6.1.4 Esquema de anotação	155
6.1.5 Processo de anotação	167
6.1.6 Processo de validação	173
6.2. Apresentação e discussão dos resultados	174
6.2.1. As relações causais no corpus	175
6.2.1.1.Estrutura e organização das relações de causa	176
6.2.1.2.Relações causais explícitas e implícitas nas notícias	182
6.2.2. Classificação dos sinais de causa	187
6.2.3. Propriedades dos eventos nas relações de causa	191
6.2.4. Relações temporais entre dois argumentos das relações causais	198
6.3. Sistematização dos resultados	207
Considerações Finais	211

Referências Bibliográficas.....	217
---------------------------------	-----

Declaração de honra

Declaro que a presente tese é de minha autoria e não foi utilizada previamente noutro curso ou unidade curricular, desta ou de outra instituição. As referências a outros autores (textos, trabalhos, ideias) respeitam escrupulosamente as regras de atribuição de autoria e encontram-se devidamente indicadas no texto e nas referências bibliográficas, de acordo com as normas de referência. Tenho consciência de que a prática de plágio e auto-plágio constitui um ilícito académico.

Declaro, ainda, que não utilizei ferramentas de inteligência artificial generativa (chatbots baseados em grandes modelos de linguagem) para realização de parte(s) da presente tese.

Porto, 23 de janeiro de 2026

Nana Yu

Agradecimentos

Os três anos desta jornada de doutoramento tiveram um significado profundo na minha vida. Como frequentemente se afirma, nada é fácil, e este percurso confirmou-o plenamente. Por essa razão, expresso o meu mais sincero agradecimento a todos quantos me apoiaram e acompanharam ao longo deste caminho, contribuindo de forma decisiva para a concretização do presente trabalho de investigação.

Expresso a minha profunda gratidão à minha orientadora, Professora Doutora Purificação Silvano, pelo acompanhamento rigoroso e dedicado ao longo desta exigente etapa académica. Constitui para mim uma honra e um motivo de grande orgulho ter sido sua orientanda, uma vez que a sua orientação foi uma presença constante e esclarecedora, iluminando o meu percurso académico. Beneficiei do exemplo de uma académica de excelência, profundamente profissional e inspiradora. Agradeço-lhe, em particular, por me ter aceitado orientar sem quaisquer condicionantes, pelas orientações sempre claras e pertinentes, pelo incentivo permanente, pela paciência, pela confiança depositada em mim e pelo apoio incondicional ao longo de todo este percurso.

Registo também o meu reconhecimento ao projeto StorySense - Uncovering deeper layers of Narrative Understanding, financiado pela Fundação para a Ciência e a Tecnologia (ref.^a 2022.09312.PTDC; DOI 10.54499/2022.09312.PTDC) pela atribuição da bolsa de investigação e pelo financiamento concedido, sem os quais a realização deste doutoramento não teria sido possível.

Manifesto igualmente o meu apreço ao Professor Doutor António Leal, pelas orientações e esclarecimentos prestados, bem como pela constante disponibilidade, paciência e confiança demonstradas. A autorização para frequentar unidades curriculares revelou-se determinante para o reforço da minha autonomia e para o aprofundamento da segurança na análise de dados em diferentes *datasets*, assim como para a minha integração no grupo de anotação e de análise de dados.

Agradeço também ao Professor Doutor Alípio Jorge, da Faculdade de Ciências da Universidade do Porto, por me receber de forma calorosa na equipa do projeto StorySense, pela generosa partilha de ensinamentos e por me ter disponibilizado um

local de trabalho no Departamento de Computação, em condições particularmente favoráveis ao desenvolvimento desta investigação.

O meu reconhecimento estende-se ainda à Professora Doutora Margarita Dimitrova, que me proporcionou conhecimentos fundamentais durante a frequência das unidades curriculares na área da Sintaxe.

Agradeço, de igual modo, ao meu orientador de mestrado, Professor Doutor António Borralho, pelo encorajamento e apoio constantes, que me incentivaram a dar mais um passo no meu percurso académico. A sua presença revelou-se especialmente significativa ao longo de todo o período de doutoramento.

Ao colega Luís Filipe Cunha, deixo o meu agradecimento pelo apoio técnico na área da computação, nomeadamente na resolução de questões relacionadas com a plataforma de anotação e na extração dos dados quantitativos utilizados neste estudo. Agradeço a disponibilidade, eficiência, atenção e amabilidade que sempre demonstrou.

Um agradecimento muito especial é devido aos amigos Sissi Azevedo, Paulo Ramos e Aline Amantes, bem como às suas famílias, pela disponibilidade constante, pelo apoio emocional, pelo carinho, pela gastronomia portuguesa e brasileira, tão importantes nos momentos mais difíceis, e pelo cuidado, atenção e acolhimento generoso que me proporcionaram, permitindo-me uma valiosa experiência de integração cultural.

Aos amigos e colegas Evelin Amorim, Nuno Guimarães, João Pedroso, Inês Cantante, Rute Rebouças, Miguel Correia, Vítor Magalhães, Samah Nassir, Aparecida Donizetti Paes, Fabrício Bitencourt, Yangyang Chen, Xinin Shi, Qinren Zuo e Zichen Zhang, agradeço a presença constante, a confiança, o incentivo e a amizade ao longo deste percurso.

Por fim, reservo uma palavra muito especial para os meus pais, Yuehu Yu e Ronghua Cao, e para o meu irmão, Hao Yu, cujo apoio incondicional e confiança permanente foram absolutamente determinantes para a concretização deste trajeto académico.

Resumo

O presente trabalho, desenvolvido no âmbito do Doutorado em Ciências da Linguagem, tem como objetivo a investigação das propriedades semânticas das relações de causalidade em notícias de natureza narrativa escritas em português europeu. A causalidade é, neste trabalho, entendida como uma relação semântica entre duas situações, em que uma funciona como argumento-razão e a outra como argumento-resultado, constituindo ambas os elementos nucleares de uma relação de Causa.

Para a caracterização da relação de Causa, foram definidas questões de investigação que orientaram todo o trabalho e incidem sobre os seguintes parâmetros: (i) a estruturação das relações causais no que diz respeito à natureza e à ordem dos argumentos de razão e de resultado; (ii) a presença ou ausência de elementos e estruturas linguísticas que sinalizam a relação causal; (iii) os sinais de causalidade mobilizados para marcar a relação de Causa; (iv) as características morfosintáticas e semânticas dos eventos principais associados a cada argumento; e (v) o papel das relações temporais na expressão da causalidade.

Com vista à concretização desta análise e após uma revisão das principais propostas teóricas sobre causalidade, tempo e aspeto, foi desenvolvido um esquema de anotação causal e temporal, LUSA-ACT, que integra contributos da norma ISO 24617 – *Semantic Annotation Framework (SemAF)*, nomeadamente da Parte 1 (*Time and Events*; ISO, 2012) e da Parte 8 (*Semantic Relations in Discourse-DR-Core*; ISO, 2016), bem como do esquema de anotação Text2Story (Silvano et al., 2021; Leal et al., 2022).

O esquema LUSA-ACT possibilitou a representação integrada da informação causal e temporal num corpus constituído por 300 notícias de natureza narrativa da Agência Lusa. Os resultados obtidos indicam que a relação causal ocorre com frequência no discurso noticioso, manifestando-se tanto em contexto intrafrásico como interfrásico e com os argumentos dispostos quer na ordem razão-resultado, quer na ordem inversa. O estudo evidencia ainda que, nas narrativas jornalísticas, as relações causais se expressam predominantemente de forma explícita, sendo menos frequentes as ocorrências implícitas; no caso das relações explícitas, os marcadores pertencem a diferentes classes gramaticais, com destaque para preposições e verbos. A análise demonstra igualmente que as eventualidades que integram os argumentos causais são

maioritariamente de natureza dinâmica e télica, sendo frequentemente realizadas por verbos e nomes e localizadas num intervalo temporal passado. Algumas das propriedades aspetuais identificadas contribuem para explicar a maior frequência de determinadas relações temporais associadas à causalidade, salientando-se, em particular, as relações de anterioridade e posterioridade.

A tese encontra-se organizada em seis capítulos. O primeiro capítulo apresenta o conceito de causalidade, numa perspetiva filosófica e linguística, e procede à distinção entre diferentes tipos de causas. O segundo capítulo revê as principais propostas de análise das relações discursivas, com especial enfoque na caracterização das relações causais. No terceiro capítulo, são analisadas as formas de marcação explícita da causalidade, com particular atenção a verbos causativos, nomes, preposições e orações com valor causal. O quarto capítulo é dedicado às categorias gramaticais de tempo e aspeto, abordando-se os seus conceitos fundamentais, bem como propostas de análise e de representação da informação temporal e aspetual. O quinto capítulo apresenta uma revisão dos principais esquemas de anotação causal e temporal, com o objetivo de situar o estado da arte neste domínio. Por fim, o sexto capítulo é dedicado ao estudo empírico da causalidade em notícias narrativas em português europeu, incluindo a descrição do estudo e a apresentação e discussão dos resultados.

A investigação desenvolvida apresenta três contribuições centrais. Por um lado, constitui um contributo para uma compreensão mais aprofundada da semântica das relações causais, permitindo identificar propriedades relevantes para a definição da relação de Causa e, conseqüentemente, para a sua inferência, quer em contextos de anotação manual, quer automática. Por outro lado, propõe um esquema de anotação, LUSA-ACT, que harmoniza, de forma integrada, as camadas causal e temporal, contribuindo para o avanço da anotação semântica. Por último, a análise desenvolvida conduziu à criação de um corpus anotado com relações causais e temporais em português europeu, que será disponibilizado em acesso aberto. Este recurso constitui um contributo relevante para a investigação linguística e computacional, colmatando uma lacuna significativa numa língua que ainda dispõe de um número limitado de recursos anotados desta natureza.

Palavras-chave: semântica; relações discursivas; causa; marcadores de causalidade; relações temporais

Abstract

This dissertation, developed within the PhD program in Language Sciences, aims to investigate the semantic properties of causal relations in narrative news texts written in European Portuguese. Causality is understood in this work as a semantic relation between two situations, in which one functions as a reason argument and the other as a result argument; together, they constitute the core components of a causal relation.

To characterize causal relations, a set of research questions was defined to guide the study, focusing on the following parameters: (i) the structuring of causal relations with respect to the nature and order of the reason and result arguments; (ii) the presence or absence of linguistic elements and structures that signal the causal relation; (iii) the causal markers used to encode the causal relation; (iv) the morphosyntactic and semantic properties of the main events associated with each argument; and (v) the role of temporal relations in the expression of causality.

To carry out this analysis, and following a review of the main theoretical approaches to causality, time, and aspect, a causal and temporal annotation scheme, LUSA-ACT, was developed. This scheme integrates contributions from the ISO 24617 standard (*Language Resource Management – Semantic Annotation Framework*), namely Part 1 (*Time and Events*; ISO, 2012) and Part 8 (*Semantic Relations in Discourse – DR-Core*; ISO, 2016), as well as from the Text2Story annotation scheme (Silvano et al., 2021; Leal et al., 2022).

The LUSA-ACT scheme enabled the integrated representation of causal and temporal information in a corpus of 300 narrative news articles from the Lusa News Agency. The results show that causal relations occur frequently in news discourse, appearing both in intra-sentential and inter-sentential contexts, with arguments arranged either in reason–result order or in the reverse order. The study further demonstrates that, in journalistic narratives, causal relations are predominantly expressed explicitly, while implicit realizations are less frequent. In the case of explicit relations, causal markers belong to different grammatical classes, most notably prepositions and verbs. The analysis also shows that the eventualities involved in causal arguments are mainly dynamic and telic in nature, are often realized by verbs and nouns, and are typically located in a past temporal interval. Some of the identified aspectual

properties help explain the higher frequency of certain temporal relations associated with causality, particularly anteriority and posteriority.

The dissertation is organized into six chapters. Chapter 1 presents the concept of causality from both philosophical and linguistic perspectives and distinguishes different types of causes. Chapter 2 reviews the main approaches to discourse relations, with particular emphasis on the characterization of causal relations. Chapter 3 is devoted to the grammatical categories of tense and aspect, addressing their core concepts as well as proposals for the analysis and representation of temporal and aspectual information. Chapter 4 analyzes explicit markers of causality, with particular attention to causative verbs, and nouns, prepositions, and clauses with causal meaning. Chapter 5 reviews the main causal and temporal annotation schemes in order to situate the state of the art in this domain. Finally, Chapter 6 presents the empirical study of causality in narrative news texts in European Portuguese, including the description of the study and the presentation and discussion of the results.

This research makes three main contributions. First, it advances a deeper understanding of the semantics of causal relations by identifying properties that are relevant to the definition of the causal relation and that may facilitate its inference, both in manual and automatic annotation contexts. Second, it proposes an annotation scheme, LUSA-ACT, that harmonizes causal and temporal layers in an integrated manner, contributing to advances in semantic annotation. Finally, the analysis conducted led to the creation of an annotated corpus containing causal and temporal relations in European Portuguese, which will be made available as an open-access resource. This corpus represents a significant contribution to linguistic and computational research, addressing an important gap in a language that still has a limited number of annotated resources of this kind.

Keywords: semantics; discourse relations; causality; causal markers; temporal relations

Índice de Figuras

FIGURA 1 - CINCO TIPOS DE ESQUEMA TEXTUAL.....	45
FIGURA 2 - ORGANIZAÇÃO DAS RELAÇÕES DISCURSIVAS	46
FIGURA 3 - CAUSA VOLITIVA E CAUSA NÃO VOLITIVA	49
FIGURA 4 - RELAÇÃO DE CAUSA - NON-VOLITIONAL CAUSE E PURPOSE.....	50
FIGURA 5 - RELAÇÃO DE CAUSA - NON-VOLITIONAL RESULT.....	53
FIGURA 6 - RELAÇÃO DE CAUSA - PURPOSE	54
FIGURA 7 - REPRESENTAÇÃO DAS FRASES EM ESTRUTURAS DE REPRESENTAÇÃO DISCURSIVA	57
FIGURA 8 - HIERARQUIA DE RELAÇÕES DISCURSIVAS	67
FIGURA 9 - OS SUBTIPOS DE CAUSA NA CLASSE CONTINGENCY	68
FIGURA 10 - REDE DE TRANSIÇÃO ASPETUAL	130
FIGURA 11 - PROCEDIMENTOS METODOLÓGICOS	152
FIGURA 12 - TEXT2STORY MULTILAYER ANNOTATION SCHEME.	157
FIGURA 13 - INTEGRAÇÃO DA RELAÇÕES DISCURSIVAS DE CAUSA	157
FIGURA 14 - METODOLOGIA DE DESENVOLVIMENTO DO ESQUEMA DE ANOTAÇÃO (<i>MATTER CYCLE</i>)	158
FIGURA 15 - ESQUEMA DE ANOTAÇÃO CAUSAL PRELIMINAR	159
FIGURA 16 - ESQUEMA LUSA DE ANOTAÇÃO CAUSAL-TEMPORAL (LUSA-ACT).....	164
FIGURA 17 - INTERFACE DA PLATAFORMA BRAT.....	168
FIGURA 18 - INTERFACE DA PLATAFORMA INCEPTION.....	169
FIGURA 19 - PROCESSO DE ANOTAÇÃO CAUSAL	170
FIGURA 20 - ANOTAÇÃO DE UM SEGMENTO TEXTUAL DO CORPUS	171
FIGURA 21 - ANOTAÇÃO DE UM SEGMENTO TEXTUAL COM UMA RELAÇÃO CAUSAL COM CONJUNÇÃO	172

Índice de Tabelas

TABELA 1 - DIFERENÇAS ENTRE RELAÇÕES DE CAUSA E RESULTADO	47
TABELA 2 - RELAÇÃO ENTRE ARGUMENTOS - NÚCLEO E SATÉLITE.....	52
TABELA 3 - RELAÇÃO DE EXPLICAÇÃO E DE RESULTADO	60
TABELA 4 - RELAÇÕES DISCURSIVAS PROPOSTAS PELA ISO 24617-8 (2016)	73
TABELA 5 - NOMES MNEMÓNICOS PARA OS PAPEIS ARGUMENTAIS DAS RELAÇÕES DISCURSIVAS ASSIMÉTRICAS	75
TABELA 6 - MAPEAMENTO DE RELAÇÃO DE CAUSA (ISO 24617-8, 2016).....	76
TABELA 7 - TESTES SINTÁTICOS PARA ANÁLISE DO COMPORTAMENTO DE DIFERENTES TIPOS DE ORAÇÕES E CONETORES.....	103
TABELA 8 - CRITÉRIOS DE DISTINÇÃO DO ESTATUTO MAIS COORDENATIVO OU SUBORDINATIVO DOS CONETORES CAUSAIS/EXPLICATIVOS.....	106
TABELA 9 - CARACTERIZAÇÃO DOS TEMPOS VERBAIS SEGUNDO REICHENBACH (1947)	116
TABELA 10 -CARACTERIZAÇÃO DOS TEMPOS VERBAIS SEGUNDO KAMP & REYLE (1993).....	119
TABELA 11 - CATEGORIAS DOS EVENTOS E ESTADOS	128
TABELA 12 - CLASSES ASPETUAIS BÁSICAS.....	132
TABELA 13 - FREQUÊNCIA DE RELAÇÕES CAUSAIS POR NOTÍCIA.....	175
TABELA 14 - CASOS DE RELAÇÕES CAUSAIS COM CONJUNÇÕES E DISJUNÇÕES.....	177
TABELA 15 - ORDEM DOS ARGUMENTOS NAS RELAÇÕES CAUSAIS	182
TABELA 16 - RELAÇÕES CAUSAIS EXPLÍCITAS E ÍMPLICAS NO CORPUS.....	183
TABELA 17 - FREQUÊNCIA E CLASSIFICAÇÃO DOS SINAIS DE CAUSA	188
TABELA 18 - ATRIBUTOS E VALORES DOS EVENTOS NAS RELAÇÕES CAUSAIS.....	192
TABELA 19 - ATRIBUTOS E VALORES DOS EVENTOS NO ARGUMENTO RAZÃO E RESULTADO.....	194
TABELA 20 - COMBINAÇÕES DOS VALORES DOS EVENTOS DOS ARGUMENTOS DE RAZÃO E RESULTADO	196
TABELA 21 - TIPOS DE RELAÇÕES TEMPORAIS NA RELAÇÕES CAUSAIS	199
TABELA 22 - RELAÇÕES TEMPORAIS NAS RELAÇÕES CAUSAIS EXPLÍCITAS	201
TABELA 23 - ORDEM DOS ARGUMENTOS NAS RELAÇÕES TEMPORAIS	202
TABELA 24 - ORDEM DOS ARGUMENTOS DAS RELAÇÕES TEMPORAIS NAS RELAÇÕES CAUSAIS EXPLÍCITAS	203
TABELA 25 - COMBINAÇÕES DOS VALORES DOS EVENTOS ASSOCIADOS A RELAÇÕES TEMPORAIS	205

Lista de abreviaturas e siglas

ARG.2	ARGUMENTO (RAZÃO)
ARG.1	ARGUMENTO (RESULTADO)
BEF.	BEFORE
AFT.....	AFTER
INCL.....	INCLUDES
ISINCL	IS_INCLUDED
SIM.	SIMULTANEOUS
OCC.....	OCCURRENCE
ST	STATE
REP	REPORTING
IACT	I_ACTION
ISTA	I_STATE
TRAN	TRANSITION
ST	STATE
PROC	PROCESS
PST	PAST
PRES	PRESENT
FUT	FUTURE
INF	INFINITIVE
GER	GERUNDIVE
PART	PARTICIPE
QI.....	QUESTÕES DE INVESTIGAÇÃO
E1.....	EVENTO1
E2.....	EVENTO2

Introdução

A causalidade constitui um dos princípios fundamentais que orientam a racionalidade humana, uma vez que o conhecimento do mundo se organiza através de redes causais que explicam, justificam, possibilitam ou impedem a ocorrência de fenómenos. Como sublinha Lopes (2004), a linguagem encontra-se profundamente marcada por referências à causalidade, tornando-a um elemento imprescindível tanto na organização da vida quotidiana como na investigação científica. De acordo com Li (2021), a análise da causalidade contribui para a compreensão dos mecanismos subjacentes aos fenómenos naturais e sociais, permitindo a sua interpretação, a previsão de eventos futuros e a tomada de decisões informadas, além de desempenhar um papel essencial na interpretação semântica do discurso.

No domínio da Semântica, a causalidade pode ser definida como uma relação que se estabelece entre dois argumentos, dos quais um corresponde à causa e o outro ao efeito. Esta relação é central para a estruturação, organização e interpretação da informação num texto, nomeadamente no discurso jornalístico. Embora diversos estudos tenham abordado esta temática (Girju et al., 2007; Prasad et al., 2008; Dunietz et al., 2017; Tan et al., 2022), a definição e a delimitação da causalidade continuam a revelar-se complexas, sobretudo quando analisadas em diferentes géneros textuais (Arrais, 1985).

Quanto ao português, a revisão da literatura evidencia ainda um número relativamente reduzido de estudos dedicados especificamente à análise da causalidade, em particular em contexto noticioso. Destacam-se, contudo, trabalhos sobre orações adverbiais causais introduzidas por conectores causais e explicativos (Lobo, 2003), sobre os aspetos semânticos, sintáticos e pragmático-discursivos da relação de causa (Lopes, 2004), sobre relações retóricas de explicação e resultado introduzidas por quando (Cunha & Silvano, 2009), sobre relações discursivas e temporais em orações causais introduzidas por porque (Silvano, 2010), bem como estudos que propõem subtipos de relações causais, como motivação, explicação, razão e causa (por exemplo, Abreu et al., 2016).

Neste enquadramento, o presente estudo pretende contribuir para o aprofundamento da investigação sobre a causalidade em português europeu, analisando de forma sistemática como as relações causais se estruturam e se organizam em notícias de carácter narrativa, provenientes da Agência Lusa.

A notícia, enquanto género jornalístico, procura tipicamente responder a seis questões fundamentais: Quem? O quê? Onde? Quando? Como? e Porquê?, sendo, por isso, expectável encontrar algumas relações causais. Segundo Gonçalves (2019), os jornalistas recorrem frequentemente à técnica da pirâmide invertida, organizando a resposta à pergunta *Porquê?* no corpo do texto, fornecendo informações sobre as causas, explicações e consequências dos acontecimentos, o que facilita a compreensão dos fenómenos sociais e naturais por parte do leitor.

Deste modo, o objetivo central que orienta este trabalho consiste em investigar como se caracteriza a causalidade em notícias da Agência Lusa em português europeu. A investigação das relações causais permite não só analisar a organização local da causalidade no discurso noticioso, mas também contribuir para a definição das propriedades da relação discursiva Causa, de modo a facilitar a sua identificação. Para além disso, este trabalho permitiu ainda a criação de um corpus anotado com relações causais e temporais em português europeu, uma língua que ainda dispõe de poucos recursos linguísticos anotados.

Tendo em conta este enquadramento, definiram-se os seguintes objetivos de investigação.

- Objetivo geral
 - Analisar as relações de causa em notícias da Agência Lusa em português europeu.
- Objetivos específicos
 - Caracterizar as relações de causa presentes em notícias, de natureza narrativa, da Agência Lusa, escritas em português europeu;
 - Identificar os elementos linguísticos envolvidos na identificação e inferência das relações causais, bem como os tipos de sinais de causa mais frequentes no corpus;

- Caracterizar as propriedades dos eventos associados aos argumentos das relações causais no corpus;
- Correlacionar as relações temporais e as relações causais identificadas no corpus.

A partir destes objetivos, formularam-se as seguintes questões de investigação (QI):

- ❖ QI1: Quais são as principais propriedades semânticas das relações de causa em notícias de natureza narrativa em português europeu?
- ❖ QI2: Qual é a configuração da organização dos argumentos das relações causais em notícias de natureza narrativa em português europeu?
- ❖ QI3: Com que frequência ocorrem relações causais explícitas e implícitas em notícias de agência em português europeu?
- ❖ QI4: Quais são os elementos linguísticos que contribuem para a identificação das relações de causa em notícias de natureza narrativa em português europeu?
- ❖ QI5: Entre as propriedades anotadas dos eventos representados nos argumentos das relações causais, que combinações de atributos ocorrem com maior frequência?
- ❖ QI6: Quais são as relações temporais estabelecidas entre os eventos principais dos argumentos nas relações causais em notícias de natureza narrativa em português europeu?

A partir dos contributos recolhidos ao longo da revisão da literatura e da reflexão sobre o estado da arte em estudos de semântica sobre causalidade, foi feito o desenho metodológico para o estudo empírico. Assim, do ponto de vista metodológico, a investigação segue sete etapas fundamentais: (i) constituição do corpus; (ii) seleção da ferramenta de anotação; (iii) definição do esquema de anotação; (iv) experimentação e validação do esquema de anotação; (v) anotação do corpus; (vi) validação dos dados anotados; (vii) análise quantitativa e qualitativa dos resultados da anotação.

A tese encontra-se organizada em seis capítulos, para além da Introdução e das Considerações Finais, refletindo uma progressão teórica e metodológica que sustenta o estudo empírico das relações de causalidade em narrativas jornalísticas em português europeu.

No Capítulo 1, procede-se à reflexão sobre o conceito de causalidade, a partir de uma perspectiva teórica abrangente, que integra contributos da filosofia Hume (1965), Davidson (1967) e Mill (1886) e da linguística (Arrais, 1985), (Shibatani, 1976) e (Soares da Silva, 1999). São discutidas diferentes tipologias de causas, incluindo distinções clássicas e contemporâneas, tais como causação permissiva e factiva (Nedyalkov & Silnitsky, 1973), causa de conteúdo e epistémica (Sweetser, 1990; Noordman & Blijzer, 2000), bem como causa direta e indireta (Galán Rodríguez, 1999), permitindo estabelecer um enquadramento concetual sólido para a análise linguística da causalidade.

O Capítulo 2 é dedicado à análise da causalidade no âmbito das principais teorias das relações discursivas. São revistas propostas fundamentais como a *Theory of Discourse Coherence* (Hobbs, 1985), a *Rhetorical Structure Theory* (Mann & Thompson, 1987) e a *Segmented Discourse Representation Theory* (Asher & Lascarides, 2003; Silvano & Cunha, 2009; Silvano, 2010) bem como extensões desta última e abordagens baseadas em corpora anotados, nomeadamente o *Penn Discourse Treebank* (Prasad et al., 2008) e a norma ISO 24617-8 (2016). Este capítulo permite situar a relação de Causa no conjunto mais amplo das relações discursivas e evidenciar os seus critérios de identificação e de inferência no discurso.

No Capítulo 3, a atenção centra-se nos mecanismos linguísticos envolvidos na inferência de causalidade no discurso. São analisadas as principais marcas linguísticas associadas à expressão da causalidade, incluindo recursos lexicais, nomes e verbos causativos, preposições, conetores e orações com valor causal, com o objetivo de compreender de que modo a causalidade é sinalizada explicitamente ou implicitamente nas línguas naturais.

O Capítulo 4 aborda as categorias gramaticais de tempo e aspeto, fundamentais para a caracterização das relações causais. São apresentadas as noções centrais destas categorias, bem como algumas das principais propostas teóricas para a sua análise como: na categoria de tempo, destacm-se os de Reichenbach (1947), Kamp & Reyle (1993) e Oliveira (2003; 2013), na de aspeto, referem-se aos de Vendler (1967), Moens &

Steedman (1988) e Cunha (2013), destacando-se modelos clássicos e desenvolvimentos posteriores.

No Capítulo 5, procede-se a uma revisão dos principais esquemas de anotação semântica de informação temporal e causal, com especial atenção a abordagens desenvolvidas no âmbito da linguística computacional e da anotação de corpora. São analisados esquemas de anotação temporal como Pustejovsky et al. (2003a), Costa & Branco (2012), Silvano et al. (2021; 2024), esquemas específicos de anotação causal com Dunietz et al. (2015; 2017), Tan et al. (2022), e propostas que integram ambas as dimensões Bethard & Martin (2008) e Rink et al. (2010) com o objetivo de identificar convergências, lacunas e desafios que informam o desenvolvimento do esquema de anotação proposto nesta tese.

Por fim, o Capítulo 6 apresenta o estudo empírico das relações causais em narrativas jornalísticas em português europeu. Este capítulo encontra-se organizado em duas partes: uma descrição detalhada do desenho do estudo, do corpus e da metodologia adotada, e a apresentação e discussão dos resultados obtidos, face ao enquadramento teórico previamente estabelecido.

Nas Considerações finais, apresentam-se as principais conclusões, limitações do estudo e perspetivas para investigações futuras.

1. Causalidade

Neste capítulo, apresenta-se o conceito de causalidade com intuito de conhecer a sua origem, pois, parafraseando Hume (1965), para uma melhor compreensão de um determinado conceito, é preciso primeiramente conhecer a sua origem e a sua relação com as diferentes áreas do saber. Segundo Li (2021), a causalidade tem sido um foco constante em várias áreas do conhecimento desde a Filosofia, a Psicologia, a Religião, a Linguística passando pela Física, pela Matemática, pela Medicina até ao Processamento de Linguagem Natural.

Desta forma, este primeiro capítulo debruça-se sobre a causalidade, especificamente, numa perspetiva filosófica e, sobretudo, linguística. Após a definição de causalidade na perspetiva filosófica e linguística, começaremos por distinguir alguns termos, que, por vezes, são usados com tendo o mesmo conceito, como causação, causalidade e causatividade, mas que, em algumas propostas, são usados com sentidos diferentes. De seguida, apresentaremos diferentes propostas de tipos de causa. Terminaremos o capítulo com uma breve síntese.

1.1. O conceito de causalidade

Desde a antiguidade, muitos estudiosos procuram conhecer a causa de todos os fenómenos tanto sociais como naturais, com o intuito de criar a lei de causalidade. De acordo com Couper-Kuhlen & Kortmann (2000, p. 10):

Causality is simply an indispensable principle in the organisation of people's everyday lives, both on the individual and the micro-group level, and, a fortiori, also on the level of societies, states, and international relations.

Por isso, a causalidade tem sido objeto de estudo em várias áreas das ciências sociais, humanas e naturais e, atualmente, também tem sido foco na área do Processamento Natural da Linguagem. Muitos estudos e diversos autores, como Shibatani (1976) e Talmy (1976), consideram que “a noção de causa, causatividade ou causação é, pois, um conceito “primitivo”, e conseqüentemente, indecomponível” (Soares da Silva, 1999, p. 551), devendo ser tomado como um conceito primitivo.

Em termos gerais, a causalidade pode ser definida como uma relação entre dois eventos, acontecimentos, ou até mesmo ideias, em que uma ação causa, ou é a causa da outra. Trata-se de uma relação de causa e efeito, na lógica de que toda a ação pressupõe uma reação. Ao nível cognitivo-concetual, Meyer (2000) defende que:

(...) it is quite conceivable that the abstract principle of causality as used in discursive reasoning should be derived from agentive causation. In this case, the cognitive experience of causality as an abstract principle would be regarded as a notion that is gained by abstraction from prototypical agentive causation (Meyer, 2000, P. 23).

Nesta conceção, a causação agentiva relaciona-se com os eventos específicos ou concretos nos quais há um agente (tipicamente mais animado) que produz um efeito através de uma determinada ação. É com base nestas experiências que o ser humano constrói o conceito de causalidade ao nível abstrato, que não depende necessariamente de agentes ou de intencionalidade. Isto significa que, em termos cognitivos, em primeiro lugar, compreendemos que há um agente que realiza uma ação que conduz a um efeito (e.g. *O Pedro partiu a janela*) e depois esse padrão é generalizado a situações sem agentes (e.g. *A chuva causou a inundação*). É assim que surge o princípio abstrato da causalidade usado no raciocínio discursivo. Desta forma, a causalidade tem como base eventos que se ligam por meios de relações de causa e efeito, sem a necessidade de especificar um agente, ou seja, não é fundamental saber quem ou o quê causou o efeito, o importante é reconhecer a relação geral de causa e efeito. A noção de causalidade é derivada das experiências concretas da causação. Este tipo de relação é evidente tanto ao nível social, desde a relação entre as pessoas até ao nível natural, em particular, a relação entre os fenómenos da natureza.

De acordo com Meyer (2000), a causação agentiva parte de um verbo agentivo e intencional, que é reconceptualizado como uma entidade abstracta expressa por uma nominalização desse verbo, que torna visível esse processo de abstração. Desta forma, a nominalização atua como uma passagem entre a causalidade experiencial (agentiva) e a causalidade concetual (abstrata).

No entanto, de acordo com Soares da Silva (1999), é importante distinguir causalidade de causação e de causatividade. De acordo com a sua proposta,

o termo *causação* (já usado na literatura filosófica e cujo correspondente do inglês "causation" é frequente utilizado na literatura linguística) para nomear uma qualquer acção de causar, o termo *causalidade* para denotar a relação entre uma causa e o seu efeito ou entre o agente e o paciente, e reservamos *causatividade* para referir a expressão linguística da *causação* (Soares da Silva, 1999, p. 550).

Há muitos estudos feitos em torno deste conceito, havendo alguns que fazem distinções motivadas por considerações metafísicas e outros por considerações linguísticas. No âmbito das distinções metafísicas, várias são as propostas que procuram caracterizar a natureza da causalidade e das causas e efeitos. Mill (1886) considera as relações causais como condicionais lógicas entre proposições, cada afirmação causal pode ser reformulada como uma implicação entre factos ou frases verdadeiras. Por outras palavras, este autor vê causas e efeitos como frases (ou factos), sendo a ligação entre eles expressa por uma condicional. Davidson (1967), contrariando Mill (1886), considera as relações causais como relações entre objetos extensionais, nomeadamente eventos, e não entre frases. Segundo Dretske (1970), o que entra verdadeiramente em relação causal não são os eventos tomados como um todo, mas certas características ou propriedades desses eventos, isto é, os aspetos que são relevantes para a produção do efeito observado. Mellor (1980) e Bennett (1988) consideram que os factos (entendidos como estados de coisas ou proposições verdadeiras) são elementos fundamentais nas relações entre causa e efeito. Por sua vez, Menzies (1989) propõe que se considere situações (conjuntos específicos de circunstâncias) como os elementos que entram em relação causal. Campbell (1990) introduz a noção de *tropos* (propriedades particularizadas de objetos ou eventos) como possíveis elementos para as relações causais. Por fim, Armstrong (1997) argumenta que os estados de coisas (combinações concretas de entidades e propriedades que fazem proposições serem verdadeiras) constituem a base subjacente das relações causais.

No campo linguístico, a causalidade tem assumido diferentes formas de concetualização e interpretação, no que diz respeito aos elementos e estruturas linguísticas que a veiculam até ao tipo de relações que são estabelecidas entre as situações. A ideia de causalidade, geralmente, tende a ser evidenciada por um verbo que expressa uma causa, incorporando "a noção de 'causa-causalidade na sua significação, sendo por isso parafraseável por 'causar uma acção ou 'causar a aquisição

de uma qualidade ou a entrada num estado” (Soares da Silva, 1999, p. 550). Contudo, a construção da causalidade é feita a partir de outros dos elementos linguísticos relevantes como os tempos, modos e aspetos verbais, os conectores discursivos, as conjunções subordinativas e coordenativas, entre outros elementos. Várias são as propostas para identificar mecanismos linguísticos que permitem a inferência das relações causais, assim como para caracterizar as causas e efeitos (Arrais, 1985; Lopes, 2004; Silvano, 2010), que apresentaremos com maior detalhe na secção 1.1.2.

1.1.1. Abordagem filosófica da causalidade

A causalidade, conhecida como causação em filosofia, é um dos temas mais centrais e complexos neste campo de saber, quer no âmbito metafísico quer no epistemológico. Na verdade, foram os filósofos os primeiros a levantar as questões como “what causal relationships in general are, what it is for one thing to cause another, or what it is for nature to obey causal laws” (Mackie, 1980, p. 1).

No entanto, atualmente considera-se a causalidade como uma temática multifacetada, não apenas pelo facto de vários filósofos terem criado e desenvolvido diferentes perspetivas em torno da mesma, como também pela mesma ser objeto de estudo de outras áreas de conhecimento social e humano.

A lógica de causalidade na filosofia está relacionada com a incessante busca de conhecimento, justificação, verdade em torno dos fenómenos sociais e naturais. Geralmente, o conhecer implica construir uma crença verdadeira, justificada na lógica de uma relação de causalidade entre o sujeito e os aspetos da realidade que o rodeiam, seguindo a crença de que tudo tem uma causa ou uma origem e nada surge do nada, tudo tem uma razão de ser.

Segundo Hume (1965), qualquer crença sobre o mundo tem de ser justificada com base na experiência. A sua perspetiva sobre a causalidade baseia-se na seguinte premissa: “The examination of the impression bestows a clearness on the idea; and the examination of the idea bestows a like clearness on all our reasoning” (p. 53). Esta premissa é sustentada em dois objetos, causa e efeito, porque, na lógica do autor, “there is nothing existing, either externally or internally, that should not be considered as a cause or effect” (1965, p. 53). Por outras palavras, “a existência de qualquer ser somente

pode ser provada mediante argumentos derivados da sua causa ou do seu efeito, e estes argumentos fundam-se inteiramente na experiência” (Hume, n.d., p. 28). De acordo com Hume, apud por Davidson (1967, p. 691), “we may define a cause to be an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second”. Esta perspectiva tende a seguir a ordem temporal dos acontecimentos, visto que o autor define a relação de causalidade como um fenómeno que resulta da ocorrência das ideias, dos acontecimentos, ou seja, eventos, que se baseia em três condições: precedência temporal da causa em relação ao efeito; contiguidade no tempo e no espaço que refere à proximidade entre os eventos que ocorrem próximos no tempo e no espaço; e conjunção constante, isto é, a repetição sistemática da sucessão entre dois eventos (A-B). Esta mesma lógica é enfatizada por Mackie (1980), ao referir que:

'an object precedent and contiguous to another, and where all the objects resembling the former are placed in a like relation of priority and contiguity to those objects that resemble the latter', and again as 'an object precedent and contiguous to another, and so united with it in the imagination, that the idea of the one determines the mind to form the idea of the other, and the impression of the one to form a more lively idea of the other' (Mackie, 1980, p. 3).

Neste sentido, o autor também define a causalidade como uma relação entre objetos ou eventos, na lógica de que a causa sempre ocorre antes do efeito numa linha de tempo, destacando que há uma contiguidade entre a causa e o efeito, ou seja, dois elementos de uma relação de causa estão próximos ou conectados no tempo e no espaço. Outro aspeto a considerar nesta perspectiva é a regularidade observável entre objetos semelhantes de causa e de efeito que apresentam uma necessidade extrínseca, isto é, um conjunto de observações regulares e consistentes entre acontecimentos, sem necessidade intrínseca.

Outra perspectiva bastante interessante em termos de causalidade é a de Mill (1886), segundo a qual a conceção da causalidade como a associação de duas ideias resultantes da experiência da sua combinação regular não é suficiente para explicar cabalmente essa relação de causalidade. Segundo este autor, os eventos envolvidos numa relação de causa podem ser classificados como simples e complexos, considerando os eventos simples como os que dependem inicialmente da observação,

e os complexos como aqueles que exigem uma investigação mais profunda, visto que a observação pura não consegue identificá-los de forma clara e confiável. Na perspectiva de Mill (1886), existem dois tipos de eventos na relação de causalidade, os que resultam da experiência humana na qual os cientistas controlam as condições e as variáveis, propondo possibilidades de manipulação para descobrir novas confirmações ou novos resultados e os que ocorrem naturalmente no qual o homem é meramente um observador e aceita as condições fornecidas pela natureza.

Para determinar as relações de causalidade, Mill (1886) apresenta quatro métodos distintos, defendendo que tanto a observação quanto a experimentação são instrumentos cruciais para a compreensão dos fenómenos naturais e para a identificação de causas e efeitos. Estes métodos visam estabelecer critérios sistemáticos que permitam reconhecer quais os fenómenos que podem ser considerados causas e quais correspondem aos respetivos efeitos. O método da concordância propõe que a causa de um fenómeno é explicada através da comparação de diferentes casos em que o mesmo ocorre, procurando-se identificar um único fator comum a todos esses casos. A premissa é de que, se em todas as situações nas quais o fenómeno é observado existe um fator em comum, então, esse fator pode ser considerado a causa do fenómeno. O segundo método, o método da diferença, consiste na comparação de dois casos que são idênticos em todos os aspetos relevantes, exceto por um único fator, ou seja, se o fenómeno ocorre num caso e não no outro, a diferença entre os dois casos é atribuída como causa ou efeito do fenómeno. O terceiro método resulta da combinação dos dois métodos anteriores, concordância e diferença, sendo usado para reforçar as conclusões, geralmente em situações na qual a experimentação direta não é possível. Este método parte da análise de situações em que um fenómeno ocorre e de outras em que não ocorre, procurando verificar se existe um fator comum presente apenas nos casos positivos, o que fortalece a inferência causal. Por fim, o quarto método é o método dos resíduos, usado para identificar causas adicionais de um fenómeno, após a explicação parcial do efeito por causas já conhecidas. Quando parte do efeito de um fenómeno é explicado por causas já conhecidas, o que resta (resíduo) é atribuído a outras causas ainda não identificadas.

Estas propostas de orientação filosófica ilustram tentativas de explicação que ajudam a humanidade a compreender os fenómenos naturais e os sociais. Considera-se o conhecimento de causa como guia da vida humana (Hitchcock, 1998) e as relações de causa e efeito como essenciais para conhecer a estrutura de conceitos e as teorias intuitivas que constituem a essência de relações humanas e dos fenómenos naturais (Keil, 1989). A abordagem filosófica dominante sobre a causalidade envolve eventos, factos e objetos em que a causa, seja um evento, um facto ou um objeto, ocorre antes do efeito, que também pode ser um evento, um facto ou um objeto (Khoo et al., 2002). Estas propostas são ainda evidência e como a causalidade é um conceito complexo e multifacetado, apresentando uma dificuldade notável em ser definido de modo preciso (Khoo et al., 2002).

1.1.2. Abordagem linguística da causalidade

Em termos gerais, os estudos linguísticos tendem a reconhecer que definir causalidade é uma tarefa bastante complexa. No entanto, segundo Shibatani (1976), “uma definição geral da expressão causativa deve ser dada em termos semânticos” (Shibatani, 1976, *apud* por Arrais, 1985, p. 42), caracterizando a “situação causativa como uma relação entre dois eventos, um evento-causa e um evento-efeito, de tal forma que a ocorrência do evento-efeito é inteiramente dependente do evento-causa” (Arrais, 1985, p. 42). Por outras palavras, a ideia da causação é baseada na observação de determinados factos que coocorrem sistematicamente no qual um indica (a causa) que determina o outro (o efeito) (Lopes, 2004; Barik et al., 2016). Arrais (1985) mostra que “em termos linguísticos, a relação causal refere-se às duas proposições, trata-se de duas proposições na estrutura semântica subjacente, uma proposição-causa e outra proposição-efeito, comumente condensadas numa única proposição com um simples verbo na estrutura de superfície” (Arrais, 1985, p. 42).

Soares da Silva (1999), que também sublinha a complexidade da noção de causalidade, salienta que o seu estudo envolve questões de natureza intra- e interlinguística, bem como a análise das estruturas dos enunciados que a exprimem. Segundo o autor, a causalidade tem sido amplamente abordada tanto no domínio da sintaxe como no da semântica, dando origem a diferentes propostas teóricas. Soares da

Soares da Silva (1999) observa que os primeiros estudos sistemáticos sobre esta temática assumiram um carácter tipológico e foram desenvolvidos no âmbito da gramática generativa, em particular da Semântica Generativa, que encontrou nos predicadores causativos um dos principais argumentos a favor das suas hipóteses teóricas. Neste enquadramento, destacam-se propostas como a decomposição lexical e a postulação de transformações pré-lexicais, destinadas a explicar a estrutura dos verbos causativos.

Em contraste com essas abordagens, Soares da Silva (1999) propõe uma abordagem cognitiva em semântica lexical, na qual a causalidade é definida como uma relação entre uma causa e o seu efeito, ou entre um agente e um paciente (Soares da Silva, 1999, p. 550). Esta definição insere-se numa perspetiva que privilegia a análise do significado lexical e das estruturas concetuais subjacentes à expressão linguística da causalidade.

Outra proposta relevante é a de Meyer (2000), que defende que as relações causais devem ser compreendidas a partir da noção de condições necessárias e suficientes, entendidas como os critérios mais adequados para justificar e explicar a ligação entre dois eventos. Nesta perspetiva, um evento pode ser concebido como condição necessária ou como condição suficiente para a ocorrência de outro, sendo a relação causal estabelecida a partir da interação entre essas condições. Meyer defende que, em muitos casos, as causas não correspondem a condições isoladamente suficientes para produzir um determinado efeito, mas antes a condições que são necessárias, embora insuficientes, para a ocorrência do resultado. Esta ideia é ilustrada por Meyer (2000) através de um excerto de Strawson (1985):

A man, say, falls down a flight of stone steps as he begins the descent. The steps are slippery and the man's mind is elsewhere. This is a sufficient explanation of his fall. But of course not every preoccupied man falls down every flight of slippery steps he descends. There is absolutely no question of our formulating or envisaging exceptionless laws, ..., to cover all such cases (Strawson, 1985, p. 131).

Neste exemplo, é possível identificar condições que, tomadas isoladamente, são insuficientes para causar a queda, como o facto de as escadas serem escorregadias ou de o homem estar distraído, uma vez que nenhuma dessas condições, por si só, garante a ocorrência do acidente. No entanto, essas mesmas condições revelam-se necessárias

no contexto específico descrito, na medida em que, se o homem tivesse sido mais cuidadoso ou se as escadas não fossem escorregadias, a queda não teria ocorrido.

Assim, estas condições constituem partes causalmente relevantes de uma mesma relação causal: são elementos insuficientes considerados isoladamente, mas necessários enquanto componentes de um conjunto mais amplo de condições que, em conjunto, conduzem ao efeito observado. A descida das escadas pelo homem integra, igualmente, esse conjunto de condições, funcionando como outra condição necessária, embora não suficiente, para a ocorrência do acidente.

Meyer (2000) sublinha ainda que o conjunto total dessas condições, considerado globalmente, é suficiente para explicar o resultado no caso em análise, uma vez que o efeito efetivamente se produziu, mas é desnecessário em termos gerais, dado que o mesmo efeito poderia ter sido causado por um conjunto diferente de condições, como, por exemplo, se o homem tivesse sido empurrado ou se estivesse embriagado e as escadas se encontrassem gastas:

The set of such conditions taken as a whole is unnecessary for the result because the same effect might also have been achieved by other causes, in an altogether different set of conditions (...) and sufficient (in that particular case) for the result because, after all, the result was brought about (Meyer, 2000, p. 13).

A causalidade no discurso, na perspectiva de Meyer (2000), é um fenómeno complexo, fluído e dinâmico, profundamente enraizado na prática comunicativa, que não pode ser reduzido a um conjunto limitado de relações retóricas. Para o autor, a causalidade deve ser entendida como uma manifestação de intenções discursivas mobilizada pelos falantes para explicar, justificar e organizar o raciocínio no texto.

Embora a causalidade seja apenas um entre vários princípios de organização textual, assume um papel particularmente relevante, uma vez que permite estruturar o raciocínio, estabelecer conexões lógicas entre eventos, ideias ou ações e, do ponto de vista do leitor, facilitar a compreensão do discurso. Com efeito, a causalidade contribui decisivamente para a construção da coerência textual, tornando mais fácil ao leitor ou ouvinte acompanhar uma narrativa ou um argumento e integrar as diferentes partes do texto num todo significativo. Assim, o princípio da causalidade organiza o texto de forma lógica, permitindo que unidades individuais contribuam para uma estrutura global coesa.

Meyer (2000) defende, contudo, que a causalidade não atua isoladamente, mas integra um conjunto mais amplo de princípios fundamentais de organização discursiva, os quais apresentam as seguintes características:

- a. They stem from basic needs arising in human communication;
- b. They therefore constitute reasons why we should need more than one clause to say what we want to say, that is, motives to utter a text rather than a single clause;
- c. They represent ways in which clauses, sentences, or larger text chunks can be relevant to each other;
- d. They are not derivable from a further common principle.

(Meyer, 2000, p. 27)

Apesar da existência destes princípios, Meyer defende que a causalidade tende a assumir um estatuto privilegiado, na medida em que satisfaz de forma particularmente clara todos os critérios enunciados. O autor sublinha que:

- a. It is a basic need in human discourse to explain, to justify, to reason about causes, conditions and consequences. The close affinity between the primitive concept of causation and human action is no coincidence. People want to know about causes, reasons and consequences because they need to act;
- b. The most natural discourse strategy of explanation or stating consequences is to add another clause;
- c. The common use of asyndetic clause connections to explicate causal relations in discourse shows that causality creates relevance on its own.

(Meyer, 2000, p. 27)

Outra forma de compreender a causalidade consiste em encará-la como um fenómeno cognitivo, enraizado no pensamento humano, que envolve um conjunto de processos racionais relacionados com a perceção da realidade no âmbito da explicação, justificação e resolução de problemas. A lógica da causalidade, entendida como o vínculo que une a causa ao efeito, desempenha um papel central e abrangente na organização do pensamento humano, assumindo igualmente uma importância epistemológica fundamental na forma como os indivíduos percecionam, interpretam, questionam a realidade que os rodeia (Lopes, 2004; Li, 2021).

No âmbito da Linguística Cognitiva, Noordman & Blijzer (2000, p. 35) defendem que a causalidade constitui um princípio estruturante da perceção e da experiência humanas, ocupando, por isso, um lugar central na cognição. Segundo os autores, a causalidade é fundamental não apenas para a representação do conhecimento humano,

mas também para outros processos cognitivos, como a previsão, a explicação e a compreensão:

is an important ordering principle of human perception and human experience, and thus a central category in human cognition. It is fundamental both to the representation of human knowledge and to other cognitive processes like predicting, explaining, and comprehending (Noordman & Blijzer, 2000, p. 35).

Nessa perspetiva, considera-se a causalidade como um fator central na organização de raciocínio, da compreensão e da explicação. Os dois autores argumentam que a interpretação das orações causais se baseia na correspondência entre a informação linguística e uma representação cognitiva do mundo real, particularmente, uma representação cognitiva da causalidade construída a partir das experiências do mundo real. Esta ideia é reforçada por Lopes (2004) ao considerar que:

a relação de causalidade [é] um aspecto essencial da cognição humana, desempenhando um papel fundamental na forma como o ser humano percebe e representa mentalmente o mundo que o rodeia. Estando na base de grande parte das acções e decisões perante as diferentes situações do quotidiano, constitui um princípio indispensável à organização da vida dos indivíduos e das sociedades e tem um papel fundamental na comunicação humana (Lopes, 2004, p. 13).

Neste quadro, o princípio de causalidade, mediado pela linguagem, permite aos indivíduos expressar as suas perceções e explicar fenómenos ou estado de coisas, seja por meio da linguagem verbal ou não verbal. Os processos de entendimento e compreensão, assim como de construção de representação das informações no discurso recorrem necessariamente à linguagem, respeitando as regras gramaticais da língua e os mecanismos próprios de cada língua.

Um outro aspeto muito importante destacado por Noordman & Blijzer (2000) diz respeito à ordenação da percepção e da experiência humanas no discurso de causalidade. Segundo os autores, o discurso de causalidade contém um conjunto de orações que funcionam como unidades de informação, organizadas de modo coerente, que permitem ao leitor/ ouvinte interpretar valores de verdade, possibilidade concentrada ou plausibilidade. As orações enquanto unidades de informação contém uma verdade reconhecida pelos falantes. Esses valores exigem implicitamente que o leitor relacione o conteúdo das orações com a representação cognitiva no mundo real (Noordman & Blijzer (2000).

Esses autores também analisam a lógica temporal da ocorrência dos enunciados que compõem o discurso de causalidade, defendendo que:

causal relations are represented as ordered cause-effect pairs where the cause temporally precedes the effect. The representation of causal relations originates in our experience in the world. We observe co-occurrences of events in the world (Noordman & Blijzer 2000, p. 36).

Esta concepção está de acordo com a proposta de Lopes (2004), segundo a qual a relação de causa “se baseia, em grande parte, na observação de que determinados factos co-ocorrem sistematicamente e de que um (*a causa*) determina o outro (*o efeito*), sendo *a causa* o estado de coisas que precede temporalmente *o efeito*” (p. 13).

De um modo geral, direta ou indiretamente, os indivíduos recorrem aos princípios da causalidade para inferir, tomar decisões e compreender a realidade que os rodeia. As línguas naturais dispõem de múltiplos recursos para exprimir relações causais, como já foi destacado (Arrais, 1985; Lopes, 2004). Contudo, a tarefa de distinguir e classificar os diferentes tipos de relações causais revela-se particularmente complexa, uma vez que, como salienta Lopes (2004), o conceito de causa não é unitário nem homogêneo, cabendo à Linguística descrever e classificar as diversas construções causais, respeitando as especificidades de cada língua.

Apresentam-se a seguir alguns casos que permitem compreender essa complexidade, considerando diferentes formas de se expressar a causalidade. Em termos gerais, a causalidade inclui evento causa e um evento efeito, tratando-se de uma relação de interdependência. Contudo, a causalidade pode ocorrer no contexto de uma frase simples, como se pode observar nos seguintes exemplos:

- (1) Eu fiz João sair.
- (2) Eu forcei o João a sair.
- (3) Eu causei a saída de João.
- (4) O barulho acordou o garoto.
- (5) O João abriu a porta.

(Arrais, 1985, p. 42)

Como se viu antes, o evento efeito encontra-se interdependente do evento causa, porém, qualquer acréscimo na frase pode comprometer a ideia de causalidade, como se pode observar nos exemplos (6) e (7).

(6) Eu fiz João sair, mas ele não saiu.

(7) João abriu a porta, mas ela não abriu.

Outros contextos relevantes são os que envolvem pedidos ou lamentos, como em (8) e (9).

(8) Eu pedi ao João que saísse.

(9) Eu lamento que o João tenha saído

As frases (8) e (9) não são consideradas orações causativas, pois na frase (8) o evento da saída de João não ocorre necessariamente em decorrência do pedido, assim como na frase (9) o evento da saída do João não dependeu da situação expressa pelo verbo matriz.

Segundo Givón (1975) (*apud* por Arrais, 1985), existe nas relações de causalidade uma estrutura que é composta por duas partes: 1ª sujeito-agente da proposição – causa e 2ª sujeito-paciente da proposição – efeito. Aquele é o que principia a ação, o que pratica a ação e este é o objeto da causa, isto é, “um considerado iniciador-agente e outro sofrendo significativa mudança de estado/posição, (sendo) dada proeminência como sujeito e objeto do verbo causativo, respetivamente” (Arrais, 1985, p. 42). Porém, esta organização que Givón (1975) generaliza a todas as relações causalidade é questionada por Arrais (1985) pelo facto de “não corresponder à realidade” (p. 42). Os exemplos agrupados em (10) e (11) são disso evidência.

(10)

a. João abriu a porta.

b. João abriu a porta com a chave.

c. A chave abriu a porta.

d. A porta abriu.

(Arrais, 1985, p. 42)

(11)

a. João matou o cachorro.

b. João matou o cachorro com um veneno.

c. Um veneno matou o cachorro.

d. O cachorro morreu.

(Arrais, 1985, p. 42)

As três primeiras frases dos (10) e (11) expressam a causalidade típica porque existe uma relação entre a proposição causa e a proposição resultado, enquanto as frases (10d) e (11d) apenas apresentam a proposição resultado, ou seja, um só argumento, o que leva a caracterizar estas proposições como não causativas. Na proposição (10b) e (11b), encontram-se três argumentos: agente (João/João) – afetado (porta/cachorro) – instrumento (chave/veneno). Nas construções (10c) e (11c), os instrumentos tornam-se sujeitos-agentes, porque quando não estiver exposto um agente, o instrumento tende a aparecer como iniciador da ação.

É de referir que as construções de (10a), (10b) e (10c), embora causativas, não são similares. Na construção (10b), o “João” é um sujeito-agente, na construção (10b), a “chave” é um instrumento-sujeito, o iniciador da ação. Assim, enquanto a frase (10 b) é possível, (12) já não é.

(12) O João abriu a porta com o vento.

(Arrais, 1985, p. 42)

Na língua portuguesa, na perspetiva de Arrais (1985), existe um esquema semântico bem definido para estabelecer relações causativas entre os argumentos de uma frase.

(I) Pc  Pe

(II) Aa(causa-Ve) Af

(III) Aa (causa-Ve) Af com-Ai

(IV) Ai (causa-Ve) Af

(V) Ac (causa-Ve) Af

(Arrais,1985, p. 43)

Sendo que *Pc* corresponde a proposição causa e *Pe* a proposição efeito; *Aa* designa o argumento agente de *Pc*; *Af* refere-se ao argumento afetado de *Pe*; *Ai* indica o instrumento de *Pc*; *Ac* é o argumento causativo de *Pc*; e *Ve* remete para o verbo de *Pe* (Arrais,1985, p. 43). Este esquema permite identificar e classificar os diferentes argumentos que participam na organização de uma relação causativa em português e, também, permite compreender como a ação é desencadeada pelo “agente”, por meio de um “instrumento”, e quem é o “afetado” por uma ação (argumento afetado), além de especificar o verbo que expressa o efeito.

Importa ainda destacar que a relação de causalidade e a construção do discurso de causa variam de uma língua para outra. De acordo com Arrais (1985), há línguas que “utilizam procedimentos morfológicos, outras sintáticos e um terceiro grupo compreende aquelas que apresentam procedimentos tanto morfológicos como sintáticos” (p. 41). Essa variação, considerando as complexidades e as especificidades de cada língua, constituiu-se como temática a ser estudada pelos diferentes linguistas, especialmente os norte-americanos a partir da segunda metade da década de 60 e durante toda a década de 70. As publicações marcantes daquela época são as de McCawley (1968), de Lakoff (1970), por exemplo, que analisaram as estruturas causativas da língua inglesa e as edições de Shibatani (1976), que desenvolveu estudos abrangendo desde fundamentos teóricos até estudos particulares das línguas inglesa, hindi, banto, húngaro, turca, mandarim.

Vários têm sido os estudos que procuram caracterizar a relação de causalidade a partir dos elementos linguísticos que incorporam/compõem o discurso, quer oral, quer escrito, isto é, a partir de elementos como verbo, conjunção, conectores, a estrutura da frase, entre outros. Estes elementos constituem pistas para se compreender as relações semânticas existentes nas situações representadas pelas frases, nas quais há um

argumento de causa que precede a ocorrência do outro, argumento de efeito. Esses argumentos podem estar numa só oração, em orações numa frase complexa, ou até em diferentes frases ou partes do texto. As relações causais, como outro tipo de relações, nomeadamente as temporais, contribuem para a coesão e coerência textuais, e são construídas como base na articulação entre processos linguísticos de natureza sintática, semântica e lexical (Silvano, 2010, p. 8).

Por exemplo, McCawley (1968) destaca o papel que os verbos causativos na construção de relações de causalidade. Segundo o autor, estes verbos incorporam o primitivo semântico “causar”, sendo um elemento não decomposto (o verbo causativo *matar*, que expressa o sentido de *causar morrer* e *causar tornar-se não vivo*).

Já Arrais (1985) propõe explicar a construção de causatividade na língua portuguesa, da seguinte forma:

- (i) descrever os padrões morfológicos e sintáticos que o português utiliza para representar a causatividade; (ii) estabelecer parâmetros sintáticos e/ou semânticos que definem a relação entre elementos em construções causativas; (iii) examinar aspectos semântico-pragmáticos da causatividade (Arrais, 1985, p. 41).

Silvano (2010) mostra como a relação de causalidade não se restringe às orações subordinadas adverbiais causais, podendo emergir noutros tipos de orações adverbiais. A autora, baseando-se na proposta de García (1999), propõe a divisão das orações adverbiais em dois grupos: “the first set subsumes the traditionally named time clauses whereas the second gathers the clauses of reason, concessive clauses, conditional clauses and clauses of purpose” (García, 1999, p. 35). Como refere Galán Rodríguez (1999), para além das orações causais, também as finais, condicionais, concessivas e consecutivas têm subjacente o princípio de causalidade, isto é, “cada una de ellas se establece entre la causa y el efecto” (p. 3599). No entanto, o estudo de Silvano (2010) também revela que mesmo as orações subordinadas temporais podem veicular valores causais.

No grupo das orações de expressão de causalidade, as orações podem focar-se na causa ou no efeito. No primeiro subgrupo estão as seguintes: a oração exprime uma causa (orações causais (13a)), uma causa colocada como hipótese (orações condicionais

(13b)) e uma causa em que o efeito lógico não aconteceu ou foi contrariado (orações concessivas (13c)).

(13)

- a. Engordas mucho porque comes abundantemente. (Causal)
- b. Si comes abundantemente, engordas mucho. (Condicional)
- c. Aunque comes abundantemente, no engordas mucho. (Concesiva)

(Galán Rodríguez, 1999, p. 3599)

Do segundo subgrupo fazem parte as que se focam no efeito que pode ser visto como um resultado ou um ponto de chegada e incluem as orações finais (14a) e as consecutivas (14b). As finais introduzem um resultado intencional ou de motivação prévia, já as consecutivas introduzem um resultado não evitável ou consequência direta da ação expressa na oração principal, como se pode observar nos exemplos que se seguem.

(14)

- a. Siempre fumo para calmar los nervios. (Final)
- b. La noche era tan oscura que favorecía a los amantes. (Consecutiva)

(Galán Rodríguez, 1999, p. 3599)

O enfoque também pode ser dado ao processo de causa-efeito, valor que é veiculado pelas orações causais e finais, que têm um valor semelhante ou compatível (15a), sendo até possível combinar as duas numa só frase (15b).

(15)

- a. Te digo esto {porque/para que} veas quiénes son tus amigos. (compatibilidade causal/Final)
- b. Há escrito sus memorias para recordar los buenos tiempos y porque necesitaba dinero. (Combinação Causal/Final)

(Galán Rodríguez, 1999, p. 3600)

No que diz respeito aos elementos linguísticos que introduzem as orações, estes podem ser analisados a partir da estrutura de causalidade proposta por Vendler (1967) e desenvolvida por Tamura et al. (2010) nos seus estudos sobre as relações de causalidade em linguística. Com base no enquadramento de Vendler, Tamura et al. (2010) propõem uma perspetiva que considera determinados elementos gramaticais relevantes para a distinção entre os participantes da relação causal, nomeadamente no que diz respeito à oposição entre factos e eventos. Segundo estes autores, a literatura tende a privilegiar a análise da causalidade em termos de factos, relegando para segundo plano o estatuto dos eventos enquanto efeitos. A partir da análise da distribuição de gerúndios, de nomes eventivos e de formas verbais, Tamura et al. (2010) concluem que, nas construções causais, as causas correspondem tipicamente a factos, enquanto os efeitos são conceptualizados como eventos. Esta distinção sustenta a afirmação de que as expressões “X é a causa de Y” e “Y é o efeito de X” não exprimem exatamente a mesma relação entre X e Y, uma vez que envolvem entidades de natureza distinta (Tamura et al., 2010).

As observações de Tamura et al. (2010) recuperam ainda análises já avançadas por Vendler (1967), que identifica expressões como “é o efeito de”, “é o resultado de” e “é a causa de” como formas linguísticas que introduzem argumentos em relações de causalidade. Vendler observa que expressões como “é o efeito de” e “é o resultado de” tendem a selecionar eventos, ao passo que a expressão “é a causa de” apresenta um estatuto híbrido, na medida em que envolve simultaneamente factos e eventos na determinação da relação causal.

1.2. Tipos de causas

Após a conceptualização desenvolvida nas secções anteriores, torna-se evidente que a causalidade, constitui um conceito complexo, objeto de múltiplas abordagens teóricas que procuram descrever e explicar a diversidade de relações causais. Com efeito, como destacam Nedyalkov & Silnitsky (1973), a identificação do tipo de significado causativo em casos concretos levanta frequentemente dificuldades, devido à multiplicidade de

fatores heterogêneos de natureza linguística e extralinguística que influenciam a configuração das relações de causalidade:

We should be aware of the fact that in a number of concrete instances we will encounter certain difficulties in defining the type of causative meaning on the basis of the given criteria, due to the large number of heterogeneous factors (linguistic and extralinguistic) which affect the nature of a causation (Nedyalkov & silnitsky (1973, p.11).

Esta complexidade e heterogeneidade de fatores subjacentes às relações causais justificam a diversidade de tipologias e classificações de causalidade, assentes em perspectivas e critérios distintos. A partir das observações de Galán Rodríguez (1999) e Silvano (2010), por exemplo, constata-se a existência de uma multiplicidade de terminologias e denominações para os diferentes tipos de causa. Esta diversidade conceitual é também destacada por Lopes (2004), segundo a qual “não há um conceito uniforme de causa”, uma vez que abrange diferentes formas de relação (Lopes, 2004, p. 17). Esta linha de pensamento é semelhante à de Tamura et al. (2010), que sublinham o carácter não homogêneo da causalidade e a consequente proliferação de denominações e classificações, em função das perspectivas teóricas e dos critérios mobilizados para a análise e descrição dos discursos que exprimem relações causais. Deste modo, e com o objetivo de tornar explícitas essas diferentes abordagens, procede-se, no subcapítulo seguinte, à apresentação de algumas das principais propostas teóricas relativas aos tipos de causas.

1.2.1. Causa singular e causa explicativa (Davidson, 1967)

Davidson (1967) defende que as relações de causalidade podem ser expressas linguisticamente por, pelo menos, dois tipos distintos de enunciados, cuja diferença é fundamental do ponto de vista lógico e ontológico: os enunciados causais singulares e as explicações causais.

Os enunciados causais singulares afirmam diretamente que um evento particular causou outro evento particular. Neste tipo de enunciado, a causalidade é concebida como uma relação extensional entre eventos, sendo estes os verdadeiros *relata* (elementos) da relação causal. Para Davidson, um enunciado do tipo “A causou B” não estabelece uma relação entre proposições ou factos linguísticos, mas sim entre acontecimentos concretos no mundo. Embora tais enunciados não expressem

explicitamente leis gerais, Davidson argumenta que a sua verdade implica a existência de uma lei causal subjacente que cubra eventos desse tipo, ainda que essa lei não seja conhecida nem formulada pelo falante.

Por contraste, as explicações causais não consistem na afirmação direta de uma relação causal entre eventos, mas antes em enunciados explanatórios que procuram justificar a ocorrência de um evento com base noutro. Nestes casos, a linguagem causal é usada para relacionar descrições ou afirmações sobre eventos, e não os eventos enquanto tais. Davidson observa que expressões como “o facto de que X ocorreu causou Y” funcionam como formas abreviadas de explicação causal, sendo mais adequadamente interpretadas como “X explica causalmente Y”. Trata-se, portanto, de relações entre enunciados ou factos descritos, e não de relações causais no sentido estrito, que permanecem, para Davidson, relações entre eventos.

Embora a proposta de Davidson (1967) tenha sido bastante influente, nem todos os autores aceitaram que apenas eventos possam funcionar como argumentos das causais. Trabalhos posteriores procuraram alargar ou reformular esta conceção, propondo outras entidades como candidatas a termos da relação causal, tais como: factos (Mellor, 1980; Bennett, 1988), propriedades ou características (Dretske, 1970), *tropos* (Campbell, 1990), estados de coisas (Armstrong, 1997), situações (Menzies, 1989) e aspectos de eventos (Paul, 2000). Entre estas propostas, a conceção de factos como elementos das relações causais tem sido uma das mais amplamente discutidas e adotadas. Ainda assim, o quadro concetual delineado por Davidson continua a constituir um ponto de referência central, ao estabelecer de forma rigorosa a distinção entre causalidade enquanto relação entre eventos e explicação causal enquanto prática discursiva.

1.2.2. Causação permissiva e causação factiva (Nedyalkov & Silnitsky, 1973)

Na proposta de Nedyalkov & Silnitsky (1973) são apresentados dois tipos de causação: causação factiva e causação permissiva, definidos em função do papel desempenhado pelo sujeito causador na produção da mudança no plano referencial.

Na causação factiva, o sujeito causador constitui a fonte primária da mudança, na medida em que realiza uma ação direta que desencadeia o evento causado. O sujeito

causado surge, neste tipo de relação, como o participante sobre o qual a ação incide, desempenhando um papel essencialmente não agentivo relativamente à ocorrência da mudança. Como referem Nedyalkov & Silnitsky (1973), neste tipo de causação “the only or primary source of changes on the referential plane is the causing subject” (p. 10). O exemplo (16) ilustra este tipo de causação, em que o sujeito causador exerce controlo direto sobre a realização do evento:

(16) Ja velel emu prijti ‘I ordered him to come.’ (factiva)

Nedyalkov & Silnitsky (1973, p. 10)

Por contraste, na causação permissiva, a fonte primária da mudança deixa de ser o sujeito causador e passa a ser o sujeito causado, sendo o papel do causador limitado a permitir ou a não impedir a ocorrência do evento. Segundo os autores, “in permissive causation the primary source of these changes is the caused subject and the role of the causing subject is reduced to permitting or obstructing the changes” (Nedyalkov & Silnitsky, 1973, p. 10). Assim, na causação permissiva, o sujeito causador não produz diretamente a mudança de estado, mas cria condições para que esta ocorra ou se abstém de a bloquear. O exemplo (17) ilustra este tipo de relação causal:

(17) Ja [ne] razrešil emu prijti ‘I [did not] allow(ed) him to come.’ (permissiva)

Nedyalkov & Silnitsky (1973, p. 10)

Deste modo, a distinção entre causação factiva e causação permissiva assenta no grau de intervenção causal do sujeito causador: enquanto na causação factiva a mudança resulta de uma ação direta do causador, na causação permissiva a mudança depende primariamente do sujeito causado, cabendo ao causador apenas um papel indireto, de autorização ou de não obstrução.

1.2.3. Causa de conteúdo e epistêmica (Sweetser, 1990; Noordman & Blijzer, 2000)

A distinção entre diferentes tipos de causa pode ser analisada a partir de uma perspectiva pragmático-discursiva, tal como proposta por Sweetser (1990). A autora defende que as relações causais expressas linguisticamente atuam em domínios conceptuais distintos, a saber: o domínio do conteúdo (conteúdo proposicional ou mundo real), o domínio epistémico (crenças, inferências e conclusões do falante) e o domínio dos atos de fala (Sweetser, 1990, p. 76). Estas diferentes interpretações são ilustradas nos exemplos seguintes:

(18)

- a. John came back because he loved her.
- b. John loved her, because he came back.
- c. What are you doing tonight, because there's a good movie on.

(Sweetser, 1990, p. 76)

Na frase (18a), a relação causal é interpretada no domínio do conteúdo, uma vez que a oração causal descreve uma ligação direta entre dois eventos do mundo real: o facto de John amar alguém constitui a causa do seu regresso. Trata-se, portanto, de uma relação causal entre estados de coisas, tal como caracterizada por Sweetser quando afirma que, no domínio do conteúdo, “the causal relation holds between real-world events or states” (Sweetser, 1990, p. 76). Em contraste, na frase (18b), a leitura de conteúdo não é plausível, uma vez que o regresso de John não pode, no mundo real, causar o amor. A interpretação adequada situa-se no domínio epistémico: o falante infere a conclusão de que John a amava a partir do facto de ele ter regressado. Como refere Sweetser, neste domínio, “the causal relation links premises to conclusions in the speaker’s reasoning” (Sweetser, 1990, p. 77). O evento mencionado funciona, assim, como base inferencial para uma conclusão expressa no discurso. Já no exemplo (18c), a oração introduzida por *because* não estabelece uma relação causal entre eventos nem entre crenças, mas fornece uma justificação para a realização de um ato de fala, correspondendo ao domínio dos atos de fala. Segundo Sweetser, neste domínio, “the

causal relation is between the speaker's reason for performing the speech act and the speech act itself" (Sweetser, 1990, p. 77).

Com base nesta proposta, Noordman & Blijzer (2000) retomam a distinção entre causa de conteúdo e causa epistémica, concentrando-se nas duas primeiras dimensões identificadas por (Sweetser, 1990). Estas duas dimensões também são denominadas relações semânticas e relações pragmáticas por (Sanders, Spooren, Noordman, 1992). A causa semântica exprime uma relação de causalidade verificada no mundo real entre dois estados de coisas, estabelecendo-se com base no conteúdo proposicional das orações e a causa pragmática exprime uma conclusão do locutor baseada numa relação causal verificada no mundo real, com base num raciocínio do locutor.

Segundo Noordman & Blijzer (2000), as relações causais de conteúdo definem-se como aquelas em que a oração causal descreve uma relação de causalidade efetiva entre eventos ou estados de coisas no mundo real: "a sentence that expresses a causal content relation describes a real-world causality between two events or states in the world" (Noordman & Blijzer, 2000, pp. 36-37). Nestes casos, as orações estão ligadas diretamente pelo seu conteúdo proposicional, como no exemplo (18a). Relativamente às relações de causa epistémica equivalentes às relações pragmáticas, correspondem a casos em que a oração causal exprime uma conclusão do falante ou do escritor, baseada numa relação causal observável no mundo, mas construída ao nível do raciocínio discursivo. Como afirmam os autores, trata-se de uma inferência "by the writer/speaker that is based on a causal relation in the world" (Noordman & Blijzer, 2000, p. 37), tal como ilustrado no exemplo (18b).

Embora haja uma distinção entre estes dois tipos de causa, Noordman & Blijzer (2000) argumentam que as relações de causa epistémica não são independentes das relações de conteúdo. Pelo contrário, as primeiras assentam nas segundas, uma vez que o raciocínio epistémico se baseia na observação de regularidades causais no mundo. Nas palavras dos autores, "an epistemic relation reflects a line of reasoning that is allowed by the co-occurrence of events or situations in the world. The justification of that reasoning is the contingency of events in the world" (Noordman & Blijzer, 2000, p. 37).

Assim, as relações epistémicas podem ser entendidas como uma projeção discursiva das relações de causalidade do mundo real.

1.2.4. Causa direta e causa indireta (Galán Rodríguez, 1999)

A proposta de Sweetser (1990) também está na base da distinção entre causa direta e causa indireta, que encontra em Galán Rodríguez (1999) uma formulação sistemática. De um modo geral, esta bipartição encontra a sua origem na distinção clássica entre causa real e causa lógica, amplamente discutida na tradição gramatical. Segundo Galán Rodríguez (1999), a causa real exprime o fundamento de uma ação ou de um estado de coisas, podendo estabelecer uma relação objetiva entre uma causa e o efeito subsequente no mundo real ou entre um motivo interno e um resultado. A causa lógica, por sua vez, tem como função justificar uma opinião, um juízo ou uma conclusão, situando-se ao nível do raciocínio do falante:

la 'causa real' expresa el fundamento de una acción, ya sea causa externa (relación entre una causa y el efecto subsecuente: La casa se ha venido abajo porque era vieja) o interna (relación entre un motivo y un resultado: Se marchó porque estaba triste). En ambos tipos la explicación que se aduce es desconocida por el interlocutor. La 'causa lógica', por su parte, justifica una opinión o un juicio previamente conocido (Puesto que somos mortales, debemos morir) (Galán Rodríguez, 1999, p. 3601)

Esta oposição foi já formulada em trabalhos anteriores, designadamente por Marín (1979), com base em Lapesa (1978), através da distinção entre causais da ação enunciada, que expõem a circunstância ou o fator que origina a ação principal, e causais do ato enunciativo, que explicam ou justificam o próprio ato de dizer. Galán Rodríguez (1999) retoma estas propostas e integra-as numa classificação semântica mais amplas das relações causais, sublinhando que conceitos como causa, razão, motivo, fundamento ou explicação têm sido frequentemente relacionados e, por vezes, confundidos, devido à ausência de delimitação clara dos seus traços semânticos e à diversidade dos esquemas sintáticos que os exprimem (Galán Rodríguez, 1999, p. 3602).

Face à complexidade e diversidade terminológica, Galán Rodríguez (1999) propõe uma classificação semântica das relações causais, distinguindo quatro tipologias básicas das relações de causa, retomadas dos estudos por Marín (1979) e Quirk et al. (1985):

1) causa-efecto o causa necesaria (es una relación objetiva): (i) Los cuerpos caen porque existe la fuerza de la gravedad;

- 2) razón-consecuencia (el hablante expresa la inferencia de la relación): (ii) Regó las flores porque estaban secas;
- 3) razón/motivo-resultado (se señala al intencionalidad de un ser animado y el resultado subsecuente): (iii) Regué las flores porque mis vecinos me lo pidieron;
- 4) presupuesto-consecuencia o causa efectiva (la causa se combina con una circunstancia —premisa— que favorece o posibilita la acción principal o conclusión: (iv) Puesto que el tiempo ha mejorado, se mantendrán las actividades previstas.

Galán Rodríguez (1999, p. 3601)

É neste enquadramento que se consolida a distinção entre causa direta, associada às relações de causa-efeito e de motivo-resultado, e causa indireta, associada às relações inferenciais, justificativas ou baseadas em pressupostos.

Contudo, como Lopes (2004) defende, esta classificação apresenta um certo grau de arbitrariedade, na medida em que alguns subtipos se distinguem sobretudo pelo papel da inferência do locutor. Segundo a autora, “os dois primeiros subtipos se diferenciam pelo facto de, no segundo, haver uma inferência por parte do locutor relativamente ao estabelecimento da relação de causa, que estaria ausente no primeiro” (p. 21). Para além disso, Lopes sublinha que o estabelecimento de uma relação causal implica quase sempre uma seleção interpretativa por parte do falante:

quando se estabelece uma relação de causa, há sempre uma escolha por parte do locutor no sentido de que ele selecciona o facto que, de entre todos os que podem estar na origem do facto causado, constituindo alternativas ou cadeias causais, considera mais pertinente numa determinada situação de comunicação. Nessa medida, a relação de causa implica quase sempre uma inferência do locutor (Lopes, 2004, p. 21).

Para além da distinção dos quatro tipos de relações causais, Galán Rodríguez (1999) propõe uma classificação semântica bipartida, distinguindo entre causais propriamente ditas (causais puras) e causais explicativas. Esta bipartição é retomada em Silvano (2010), que observa a existência de um consenso alargado na literatura quanto à distinção entre orações de razão direta, também designadas como causais de conteúdo ou *de re*, e orações de causa indireta, igualmente denominadas causais da enunciação ou *de dicto*, em consonância com autores como Quirk et al. (1985), Galán Rodríguez (1999), Lobo (2003), Brito (2003) e Lopes (2004) e Silvano (2010).

As causais propriamente ditas correspondem às relações de causa direta, na medida em que estabelecem uma ligação causal entre dois estados de coisas no mundo

real. Segundo Galán Rodríguez (1999), neste tipo de construcións “se establece una conexión no consabida entre las oraciones A y B”, independentemente de os factos envolvidos serem ou non presupostos (Galán Rodríguez, 1999, p. 3602). Estas relacións poden asumir a forma de causa–efeito ou de motivo–resultado, como ilustrado nos exemplos seguintes em (19).

(19)

- a. Las flores se han marchitado porque no tenían agua suficiente.
- b. Te he comprado un regalo porque era tu cumpleaños.

(Galán Rodríguez, 1999, p. 3602)

No exemplo (19a), a falta de água é a causa directa do efecto observado (as flores terem murchado). Já no exemplo (19b), o motivo de ter comprado o presente é o facto de a pessoa estar a fazer anos, estabelecendo-se uma relação de motivação-resultado, em que o sujeito da ação terá de ser uma entidade animada, dotada de volição, que participa ativamente no efecto da relação de causalidade.

No segundo grupo, o das causais explicativas, Galán Rodríguez (1999) distingue entre causais explicativas próprias e causais explicativas hipotéticas:

presentan un hecho (B) que, a juicio del hablante, puede ser una explicación razonable o una justificación apropiada del hecho A, bien porque se conozca de antemano al relación A-B, bien porque el hecho B sea tal que favorece o propicia el hecho A. Aunque en ambos casos se trate de una explicación, es necesario diferenciar en este grupo entre causales explicativas propias' (circunstancias favorables o habituales —conocidas o presupuestas— de una acción) y 'causales hipotéticas' (deducciones que realiza un interlocutor basándose en su conocimiento de los hechos) (Galán Rodríguez, 1999, p. 3602).

Assim, as causais explicativas próprias remetem para circunstâncias favoráveis ou habituais conhecidas ou pressupostas, e causais explicativas hipotéticas exprimem deduções realizadas pelo locutor com base no seu conhecimento dos factos (Galán Rodríguez, 1999, p. 3602). Esta distinção é retomada por Lopes (2004), que associa as primeiras à explicação de circunstâncias e as segundas à causa da enunciação ou *de dicto* (Lopes, 2004, p. 24). Os exemplos em (20) ilustram a distinção proposta por Galán Rodríguez (1999):

(20)

- a. Ya que está todo visto, propongo que nos vayamos. (Circunstâncias)
- b. Como hace frío, las carreteras están heladas. (Causa habitual)
- c. No se ha ido, porque tiene ahí la cartera. (Causa hipotética)

(Galán Rodríguez, 1999, p. 3602)

A diferença entre as construções causais puras e as causais explicativas está no facto de as primeiras apresentarem conexão causal, isto é, conjunções e articuladores de causa, independentemente de os factos conectados serem pressupostos, ou não, e as causais explicativas apresentarem a relação causal como expectativas dentro da normalidade ou do que é habitual, assentando numa pressuposição (Rodríguez, 1999).

Lobo (2003, p.69) subscreve esta proposta e distingue entre orações causais propriamente ditas, em que a subordinada exprime a causa do estado de coisas expresso na oração principal, e orações causais explicativas, interpretadas como causa da enunciação. Nos primeiros casos, a relação causal é objetiva, nos segundos casos, trata-se de uma causalidade indireta, associada ao plano do discurso e da enunciação (Lobo, 2003, pp. 69-70), como ilustrado nos exemplos seguintes:

(21)

- a. As flores murcharam porque ninguém as regou.
- b. Ela regou as flores porque estavam secas.

(22)

- a. O João está cá, porque já vi o carro dele.
- b. Fecha a janela, porque está frio.

Deste modo, a distinção entre causa direta e causa indireta, embora com designações diferentes, revela-se transversal às diferentes abordagens teóricas, permitindo integrar de forma coerente perspetivas semânticas, pragmáticas e discursivas da causalidade, tal como desenvolvidas por Sweetser (1990), Galán Rodríguez (1999) e Noordman & Blijzer (2000).

1.3. Conclusão

Neste capítulo, apresentamos a definição do conceito de causalidade e a caracterização de diferentes tipos de causa, evidenciando a diversidade de aceções que este conceito assume em distintas áreas do conhecimento, como a Linguística e a Filosofia, bem como no âmbito de cada uma dessas áreas.

A análise realizada evidencia, igualmente, o papel fundamental da causalidade na experiência humana e a relevância do seu estudo, em particular no domínio da Semântica. Trata-se de um conceito complexo, mas onnipresente nos discursos orais e escritos, que refletem as relações causais subjacentes à organização da realidade. Neste sentido, a causalidade revela-se especialmente adequada a uma abordagem semântica, como a que se propõe desenvolver com este projeto de investigação.

2. A Causalidade nas teorias das relações discursivas

A comunicação constitui a base de qualquer interação humana, envolvendo a necessidade constante de expressar, partilhar e transmitir ideias, informações, emoções e sentimentos. Neste processo, de forma consciente ou intuitiva, o emissor organiza o seu discurso de modo a garantir que este seja coeso e coerente, de forma a possibilitar a sua interpretação pelo interlocutor. A construção discursiva não resulta, assim, de uma simples justaposição de enunciados, mas de uma organização estruturada que visa a produção de sentido, num processo dinâmico de comunicação e negociação entre interlocutores, no qual se constroem textos que integram múltiplos elementos linguísticos e conceptuais. Trata-se de um fenómeno multifacetado, caracterizado por um processo contínuo de construção e reconstrução de sentidos. Um texto, oral ou escrito, é constituído por um conjunto de ideias e de orações que podem ser entendidas como unidades discursivas de menor dimensão, frequentemente designadas como microdiscursos, cuja articulação dá origem a uma estrutura discursiva de nível superior, o macrodiscurso.

Esta conceção encontra eco em diferentes propostas teóricas. Para Couper-Kuhlen & Kortmann (2000), o texto pode ser descrito como *“a set of relations, conceptual in nature but instantiated linguistically, which can be said to hold typically between clauses or sequences of clauses in discourse”* (p. 1). De forma convergente, Hobbs (1985) concebe o discurso como uma estrutura organizada por relações de coerência, enquanto Mann & Thompson (1987) o descrevem como uma rede de relações de sentido estabelecidas entre segmentos discursivos. Em todos estes enquadramentos, o discurso é entendido como uma composição que articula unidades linguísticas ligadas por relações de sentido.

A construção de um discurso coeso e coerente implica, assim, o estabelecimento de relações de sentido entre as diferentes orações, frases ou partes que o constituem. Estas relações são habitualmente designadas, no âmbito da Linguística, por relações discursivas ou relações retóricas e constituem um relevante objeto de estudo em Semântica. São estas relações que permitem ao intérprete compreender como os diferentes segmentos do texto se articulam e contribuem para a construção do

significado global. Servem ainda estas relações discursivas para formalizar a representação do sentido que é construído no discurso, pelo que diversas teorias as têm integrado nos seus pressupostos teóricos.

A literatura apresenta uma diversidade significativa de denominações e enquadramentos teóricos para estas relações, tanto no domínio da Linguística como noutras áreas afins. Entre as designações propostas encontram-se, por exemplo, *combinations of predications* (Longacre, 1976), *rhetorical predicates* (Grimes, 1975), *sequiturity relations* (Fillmore, 1974) e *logical-semantic connectives* (Crothers, 1979; Silvano, 2010).

Nesta tese, adota-se a denominação relação discursiva, entendida como a articulação estabelecida entre duas unidades de sentido no discurso. Esta opção está em consonância com a definição proposta pela norma ISO 24617-8 (2016), segundo a qual as relações discursivas se estabelecem entre duas situações, eventualidades, factos, proposições, condições, crenças ou atos de diálogo, podendo ser realizadas por expressões linguísticas simples ou complexas, como orações, nominalizações, enunciados ou segmentos discursivos constituídos por várias frases num texto escrito ou por sequências de enunciados num discurso oral (ISO 24617-8, 2016, p. 2).

Nos últimos anos, a causalidade tem sido formalizada no âmbito de teorias linguísticas que integram diferentes tipos de relações discursivas. O presente capítulo incide sobre algumas dessas propostas fundamentais, apresentando uma revisão das abordagens desenvolvidas por Hobbs (1985), Mann & Thompson (1987), Asher & Lascarides (2003), Cunha & Silvano (2009), Silvano (2010), Prasad et al. (2008) e pela ISO 24617-8 (2016). Cada uma destas teorias propõe uma concetualização específica das relações discursivas, bem como dos seus tipos e dos mecanismos linguísticos responsáveis pela sua codificação.

2.1. *Theory of Discourse Coherence* (Hobbs, 1985)

Hobbs (1985) defende que a interpretação de um discurso resulta de um processo inferencial baseado no conhecimento, o qual se encontra em constante construção e atualização ao longo da compreensão textual. Como sintetiza Silvano (2010, p. 171), Hobbs “assumes that in the interpretation of discourse the accumulated knowledge is

used to build a theory of what is being said at the moment. Accordingly, a theory of coherence relations must be part of a knowledge-based theory of discourse interpretation". Assim, a interpretação discursiva não depende exclusivamente da informação linguística explícita, mas da interação entre o conteúdo do texto e o conhecimento previamente disponível ao leitor/ ouvinte.

Para Hobbs (1985), um texto constitui uma unidade estruturada, composta por partes interligadas por diferentes tipos de relações, tais como exemplificação, paralelismo, contraste, sequência temporal, explicação, entre outras. Estas relações, designadas pelo autor como relações de coerência, ligam segmentos discursivos adjacentes e permitem que o texto seja interpretado como um todo coerente. É precisamente a identificação destas relações que torna possível a interpretação do discurso, na medida em que elas explicitam como os diferentes segmentos contribuem para a construção global do sentido.

O autor salienta, contudo, que o processo de compreensão de um texto é complexo e exige a mobilização de diversos tipos de conhecimento por parte do leitor, e, implicitamente, do escritor. Para lidar com essa complexidade e explicar como se processa a interpretação discursiva, Hobbs (1985) identifica seis componentes fundamentais envolvidos nesse processo. Em primeiro lugar, a notação lógica ou representação do conhecimento (*logic notation or knowledge representation*), que fornece uma estrutura formal para traduzir o conteúdo discursivo numa forma lógica adequada à realização de inferências. Em segundo lugar, a tradução sintática e semântica (*syntax and semantic translation*), responsável por converter frases da língua natural em representações lógicas que permitam analisar as suas estruturas gramaticais e os significados das palavras e das frases. Em terceiro lugar, a codificação do conhecimento (*knowledge encoding*), que diz respeito à incorporação de informações sobre o mundo e sobre a linguagem numa base de conhecimento necessária à interpretação do discurso. Em quarto lugar, os mecanismos dedutivos (*deductive mechanisms*), entendidos como o conjunto de regras e procedimentos utilizados para realizar inferências lógicas a partir do conhecimento armazenado e da informação fornecida pelo texto. Em quinto lugar, as operações discursivas ou a especificação das

interpretações possíveis (*discourse operations or specification of possible interpretations*), que visam resolver problemas como a correferência e a ambiguidade. Por fim, a especificação da melhor interpretação (*specification of the best interpretation*), que consiste na seleção da interpretação mais simples e mais relevante entre as alternativas disponíveis (Hobbs, 1985).

No que diz respeito à classificação das tipologias das relações de coerência, Hobbs (1985) parte de uma caracterização geral da situação comunicativa em que ocorre o discurso entre um falante e um ouvinte. Segundo o autor, o falante pretende transmitir uma mensagem, essa mensagem tem um determinado objetivo, deve ser relacionada com o conhecimento prévio do ouvinte e apresentada de modo a facilitar o processo de compreensão. É a partir destas considerações que Hobbs propõe a existência de quatro grandes classes de relações de coerência.

A partir desta perspectiva, a interação discursiva não se reduz a um simples processo de emissão e recepção de enunciados. Pelo contrário, envolve uma coordenação entre falante e ouvinte, na qual o falante procura orientar o discurso de forma a permitir que o ouvinte reconheça o objetivo comunicativo subjacente. Para minimizar eventuais dificuldades de compreensão, torna-se essencial que o discurso estabeleça ligações claras com o contexto, com o conhecimento do ouvinte e com a situação comunicativa em que ocorre, fatores que contribuem decisivamente para a construção da coerência discursiva.

Deste modo, o discurso é entendido como um processo dinâmico, resultante da interação contínua entre falante e ouvinte. O papel do falante não se limita à transmissão de informação, mas inclui também a tarefa de criar ligações inferenciais entre o que é dito e o que o ouvinte já sabe, procurando reduzir o esforço de interpretação necessário à compreensão. As relações de coerência desempenham, assim, um papel central, na medida em que orientam as inferências que o ouvinte deve realizar e contribuem para a interpretação global do discurso.

As quatro grandes classes de relações de coerência propostas por Hobbs (1985) são as seguintes: (1) *Occasion relations*, (2) *Evaluation relations*, (3) *the ones that evoke the listener's prior knowledge* and (4) *the Expansion relations*.

Quanto à primeira classe, a relação de Ocasão (*occasion relation*) baseia-se em dois segmentos consecutivos do discurso (considerando S0 como o segmento atual, que pode ser uma oração ou excerto, e S1 como o segmento imediatamente anterior). Esta relação ocorre quando se pode inferir uma mudança de estado a partir da asserção de um segmento, de modo que o estado inicial ou final dessa mudança é inferido a partir do outro segmento. Por outras palavras, um dos enunciados descreve um estado cujo desfecho é realizado no enunciado seguinte, ou vice-versa. Essa situação é ilustrada pela sequência de instruções em (23):

(23)

- a. Walk out the door of this building
- b. Turn left.
- c. Go to the corner.

(Hobbs, 1985, p. 10)

A segunda classe de relações de coerência é a relação de avaliação (*Evaluation relation*), usada para verificar se o objetivo do discurso está a ser atingido. A relação de Avaliação estabelece-se quando, a partir de um segmento S1, se infere que o segmento S0 é um passo num plano para alcançar algum objetivo do discurso, ou, inversamente, quando S1 é inferido (com base em S0) como parte de um plano que visa a cumprir o objetivo pretendido. Em geral, esse objetivo corresponde ao propósito conversacional, que pode até mesmo ser simplesmente entreter o interlocutor. Por exemplo, no enunciado (24), o falante introduz uma história engraçada, indicando que deseja entreter o ouvinte:

(24) The funniest thing happened to me.

(Hobbs, 1985, p. 12)

A terceira classe liga um segmento do discurso aos conhecimentos prévios do ouvinte. Dentro dessa categoria incluem-se principalmente a Relação de Enquadramento (*Background*) e a Relação de Explicação (*Explanation*). A relação de

Enquadramento ocorre quando um enunciado estabelece um contexto ou cenário em que outro enunciado se insere. Formalmente, infere-se a descrição de um sistema de entidades e relações a partir de S0; em seguida, infere-se de S1 que uma entidade é colocada ou se move nesse sistema, usando-o como contexto para a ação ou informação apresentada. O exemplo a seguir em (25) ilustra essa relação de enquadramento, no qual o primeiro enunciado estabelece o cenário temporal e espacial, e o segundo introduz um evento dentro desse contexto:

(25) And one Sunday morning about ohhhh five o'clock in the morning I sat down in the Grand- no no, not in the Grand Central, in the Penn Station, and while I was sitting there a young cat came up to me...

(Hobbs, 1985, p. 12)

Já a relação de Explicação ocorre quando um estado ou evento descrito num segmento fornece a causa ou a razão para a situação apresentada noutro segmento, ou seja, infere-se que o acontecimento afirmado em S1 causou (ou poderia causar) o evento ou estado afirmado em S0. A relação de Explicação vincula, portanto, um enunciado que relata um facto ao enunciado anterior ou subsequente que justifica ou explica por que aquele facto ocorreu. Os exemplos em (26) ilustram essa relação: o enunciado (26a) descreve uma situação cujo motivo é esclarecido pelos enunciados seguintes (26b) e (26c):

(26)

- a. He was in a foul humor.
- b. He hadn't slept well that night.
- c. His electric blanket hadn't worked.

(Hobbs, 1985, p. 12)

A quarta e última grande classe de relações de coerência proposta por Hobbs (1985) envolve os casos em que um segmento expande ou amplia o conteúdo do outro,

facilitando a compreensão por meio de inferências adicionais. Essas relações inferenciais entre segmentos de texto se manifestam-se frequentemente com o recurso a conectores discursivos (conectores inferenciais) e servem para desenvolver ou contrastar ideias, tornando o discurso mais claro para o ouvinte/leitor. De acordo com a natureza dessas expansões, as relações de expansão subdividem-se em dois grupos: (i) relações positivas, como Paralelo (*Parallel*), Elaboração (*Elaboration*), Generalização (*Generalization*) e Exemplificação (*Exemplification*), nas quais um enunciado acrescenta informação, detalha, generaliza ou exemplifica o conteúdo do outro; e (ii) relações negativas, como Contraste (*Contrast*) e Expectativa Frustrada (*Violated Expectation*), nas quais um enunciado se opõe ou contradiz de alguma forma o conteúdo esperado a partir do outro.

2.2. *Rhetorical Structure Theory* (Mann & Thompson, 1987)

A teoria de Mann & Thompson (1987) propõe uma abordagem das relações discursivas centrada na articulação entre sentido e estrutura do discurso, partindo da ideia de que um texto é organizado por “relations that hold between parts of the text” (p. 1). Esta proposta, conhecida como *Rhetorical Structure Theory* (RST), tem sido muito utilizada como instrumento de análise textual, na medida em que oferece um conjunto de características particularmente adequadas ao estudo do discurso. Como referem os autores, esta proposta:

provides a combination of features that has turned out to be useful in several kinds of discourse studies. It identifies hierarchical structure in text. It describes the relations between text parts in functional terms, identifying both the transition point of a relation and the extent of the items related. It provides comprehensive analyses rather than selective commentary. It is insensitive to text size, and has been applied to a wide variety of sizes of text (Mann & Thompson, 1987, p. 2).

Do ponto de vista destes autores, uma relação discursiva estabelece-se entre dois segmentos de texto não sobrepostos, designados por núcleo (N) e satélite (S). O núcleo corresponde à parte central do enunciado, enquanto o satélite desempenha uma função de apoio relativamente ao núcleo. Cada relação discursiva é caracterizada por quatro componentes fundamentais: (i) restrições aplicáveis ao núcleo; (ii) restrições aplicáveis ao satélite; (iii) restrições relativas à combinação entre núcleo e satélite; e (iv)

o efeito comunicativo produzido pela relação. Estes elementos são essenciais para a análise estrutural do texto, uma vez que permitem descrever de que modo uma ideia principal se articula com ideias secundárias que a sustentam, esclarecem ou desenvolvem.

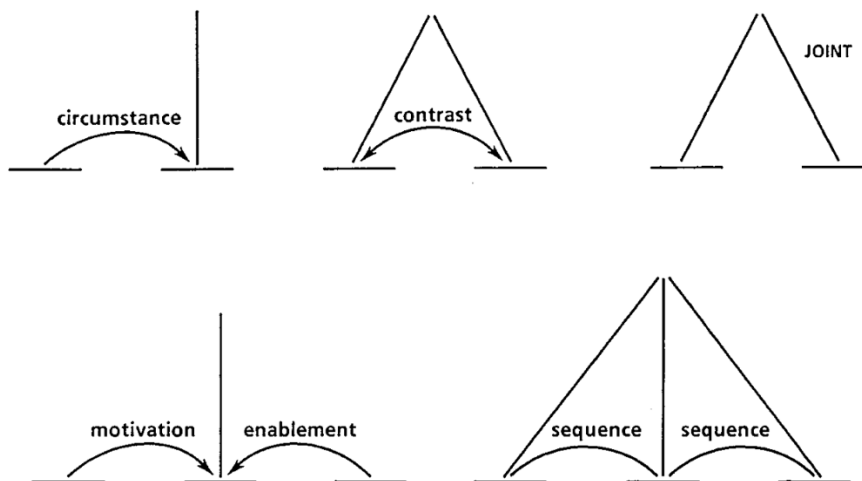
Nesta perspectiva, o texto é entendido como um discurso estruturado por um conjunto de relações de sentido que se estabelecem entre os seus diferentes segmentos. Núcleo e satélite constituem, assim, unidades dotadas de significado próprio, cuja combinação contribui para a construção do significado global do texto. Destas combinações resultam diferentes tipos de relações discursivas, tais como Evidência, Justificação, Antítese, Concessão, Circunstância, Solução, Elaboração, Enquadramento, Capacidade, Causa, Resultado, Propósito, Condição, Oposição, Interpretação, Avaliação, Atualização, Sequência e Contraste (Mann & Thompson, 1987). A análise e a compreensão de um texto implicam, portanto, a identificação destas relações e a interpretação do papel que cada segmento desempenha na organização do discurso.

Outro aspeto relevante desta teoria é o conceito de esquema textual, entendido como um padrão abstrato de organização do texto. Segundo Mann & Thompson (1987), os esquemas correspondem a;

structural constituency arrangements of text. They are abstract patterns consisting of a small number of constituent text spans, a specification of the relations between them, and a specification of how certain spans (nuclei) are related to the whole collection. They are thus loosely analogous to grammar rules (Mann & Thompson, 1987, p. 5).

Os esquemas textuais especificam, assim, as formas segundo as quais diferentes segmentos de texto podem coocorrer e articular-se, desempenhando um papel determinante na estruturação global do discurso. A partir desta perspectiva, os autores propõem cinco tipologias de esquemas de estruturação textual, que evidenciam diferentes modos recorrentes de organização das relações discursivas no texto (cf. Figura 1).

Figura 1 - Cinco tipos de esquema textual



Fonte: (Mann & Thompson, 1987, p. 7)

As curvas representam as relações que se mantêm entre os segmentos discursivos, enquanto as linhas retas indicam a identificação dos intervalos nucleares. Deste modo, distinguem-se diferentes esquemas: o esquema de *Circumstance*, no qual se estabelece uma única relação entre um núcleo e um satélite; o de *Contrast*, que envolve dois núcleos; o de *Joint*, que admite um número indefinido de núcleos; o de *Motivation-Enablement*, que inclui dois núcleos e um satélite; e o de *Sequence-Sequence*, que comporta um número indefinido de núcleos, um por cada elemento da sequência, ligados por uma relação de sucessão entre núcleos adjacentes (Mann & Thompson, 1987).

A utilização destes esquemas é regulada por três convenções fundamentais:

1. unordered spans: The schemas do not constrain the order of nucleus or satellites in the text span in which the schema is applied;
2. optional relations: For multi-relation schemas, all individual relations are optional, but at least one of the relations must hold;
3. repeated relations: A relation that is part of a schema can be applied any number of times in the application of that schema.

(Mann & Thompson, 1987, p. 6)

No processo de análise da estrutura de um texto, o primeiro passo, após uma leitura inicial de reconhecimento, consiste na sua divisão em unidades, isto é, em segmentos de ideias. Esta segmentação baseia-se numa classificação considerada teoricamente neutra, devendo cada unidade possuir integridade funcional própria. As unidades correspondem, em regra, a orações, entendidas como portadoras das ideias expressas. Existem, contudo, exceções: os sujeitos, os complementos causais e as orações relativas restritivas são considerados partes integrantes das orações em que ocorrem, não constituindo unidades autónomas (Mann & Thompson, 1987).

Assim, na análise estrutural do texto aplicam-se esquemas que procuram assegurar princípios como a completude, a conectividade, a exclusividade e a adjacência. A organização e a estrutura das unidades textuais tendem, por conseguinte, a assumir uma configuração hierárquica semelhante à de uma árvore, refletindo relações de dependência entre unidades discursivas no interior do texto. Neste enquadramento, a teoria propõe dois grandes grupos de definições de relações, organizados de acordo com o tipo específico e com critérios de semelhança, sendo que cada grupo apresenta características comuns, conforme ilustrado na Figura 2 que se segue.

Figura 2 - Organização das relações discursivas

Table 1: Organization of the Relation Definitions

Circumstance	Antithesis and Concession
Solutionhood	Antithesis
Elaboration	Concession
Background	Condition and Otherwise
Enablement and Motivation	Condition
Enablement	Otherwise
Motivation	Interpretation and Evaluation
Evidence and Justify	Interpretation
Evidence	Evaluation
Justify	Restatement and Summary
Relations of Cause	Restatement
Volitional Cause	Summary
Non-Volitional Cause	Other Relations
Volitional Result	Sequence
Non-Volitional Result	Contrast
Purpose	

Fonte: (Mann & Thompson, 1987, p. 9)

Entre as diferentes relações propostas por Mann & Thompson (1987), destaca-se a definição de relações de causa, que constitui o foco central desta investigação. A relação de causa é um dos tipos de relações discursivas mais complexo, na medida em que abrange situações muito diversas, incluindo tanto resultados intencionalmente visados por uma ação como casos de causalidade que não envolvem qualquer intenção, como a causalidade física. Por esse motivo, os autores consideram que se trata de uma relação “hard to include both situations that are intended outcomes of some action and causation that does not involve intended outcomes, such as physical causation” (Mann & Thompson, 1987, p. 57), propondo, assim, uma subdivisão interna da relação de causa. Esta subdivisão baseia-se, por um lado, na nuclearidade da situação causadora e, por outro, no caráter volitivo ou não volitivo da relação. A Tabela 1 apresenta as diferenças entre as relações de causa e de resultado, cruzando estes dois critérios.

Tabela 1 - Diferenças entre relações de causa e resultado

Differences in volition	Differences in nuclearity	
	Volitional Cause	Volitional Result
	Non-Volitional Cause	Non-Volitional Result

Fonte: adaptado de (Mann & Thompson, 1987, p. 57)

Segundo os autores, a distinção entre as relações de Causa e Resultado assenta na posição nuclear da situação causadora na estrutura discursiva. Quando a situação causadora ocupa a posição de satélite, a relação é classificada como Causa, podendo ser volitiva ou não volitiva. Quando, pelo contrário, a situação causadora é nuclear e a situação causada assume um papel menos central, a relação é classificada como Resultado, igualmente distinguindo entre volitivas e não volitivas. Dentro de cada par Causa/Resultado, a distinção entre volitivo e não volitivo depende da presença ou ausência de uma ação intencional. Os autores sublinham ainda que a relação de Propósito é, por definição, neutra quanto à volição, incluindo tanto casos volitivos como não volitivos (Mann & Thompson, 1987, p. 57).

A causa volitiva envolve a ação de um agente, tipicamente humano, que controla uma ação da qual resulta a situação nuclear. Essa ação é realizada porque o agente prefere o resultado ou, em alguns casos, a própria ação. Já a causa não volitiva corresponde a situações de consequencialidade que não implicam uma escolha ou intenção por parte de um agente, constituindo o que os autores designam como um resíduo de causalidade sem resultado escolhido (Mann & Thompson, 1987, p. 57).

No caso específico da relação de Causa Volitiva, a estrutura envolve duas dimensões principais: o núcleo (N), que contém uma ação volitiva ou uma situação resultante de uma ação volitiva, e o satélite (S), que pode não apresentar, isoladamente, uma ação explícita. Contudo, quando combinados, os segmentos discursivos permitem inferir que o conteúdo do satélite funciona como a causa da vontade do agente expressa no núcleo. Sem a presença do satélite, o efeito ou resultado descrito no núcleo poderia não ser interpretado como motivado ou poderia não ser possível identificar a motivação específica da ação. Nesta relação, o núcleo é mais central para os objetivos comunicativos do locutor do que o satélite.

Esta configuração pode ser observada nos exemplos seguintes:

(27)

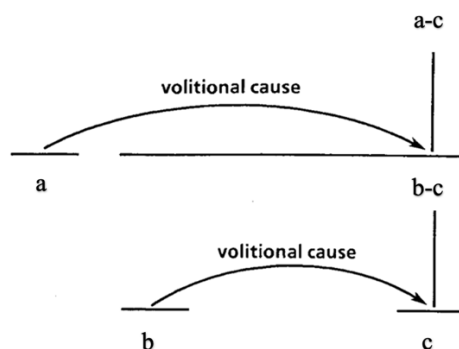
- a. Writing has almost become impossible.
- b. We had the typewriter serviced.
- c. And I may learn to type decently after all these years.

(Mann & Thompson, 1987, p. 58)

Consideradas isoladamente, as unidades (27a), (27b) e (27c) não estabelecem, por si só, uma relação causal explícita, funcionando como segmentos discursivos independentes. No entanto, quando articuladas, permitem a construção de uma relação de causa. A unidade (27a) fornece o contexto causal que justifica a ação volitiva expressa em (27b) e o resultado esperado em (27c). Embora esta relação possa aproximar-se da noção de motivação, distingue-se dela na medida em que a motivação tende a incitar o leitor ou interlocutor à realização de uma ação, enquanto a causa volitiva diz respeito à determinação da vontade do agente relativamente à ação descrita no discurso. Uma

relação semelhante pode ser identificada entre as unidades (27b) e (27c), sendo (27b) interpretada como causa volitiva do resultado expresso em (27c), uma vez que a reparação da máquina de escrever possibilita a aprendizagem de uma escrita mais competente. Esta articulação é ilustrada de forma esquemática na Figura 3.

Figura 3 - causa volitiva e causa não volitiva



Fonte: (Mann & Thompson, 1987, p. 59)

No que diz respeito à causa não volitiva, os autores apresentam a seguinte definição. Nesta relação, o núcleo (N) exprime uma situação que não resulta de uma ação volitiva e o satélite (S) não envolve igualmente qualquer ação intencional nem impõe restrições à ação. No entanto, quando considerados em conjunto, os segmentos permitem inferir que o conteúdo do satélite apresenta uma situação que, por meios distintos da motivação de uma ação volitiva, causou a situação expressa no núcleo. Assim, sem a presença do satélite, o resultado poderia não permitir identificar a causa específica da situação descrita no núcleo. Ainda que o satélite contribua para a interpretação causal, o núcleo mantém-se mais central para os objetivos comunicativos do locutor na combinação N-S. Deste ponto de vista, o resultado reconhece a situação apresentada no satélite como explicativa da situação nuclear. Esta relação pode ser ilustrada através de um excerto de um anúncio de relações públicas de uma empresa australiana do setor mineiro, no qual se justifica a exploração de recursos naturais, destacando os seus benefícios económicos para a população australiana (cf. 28):

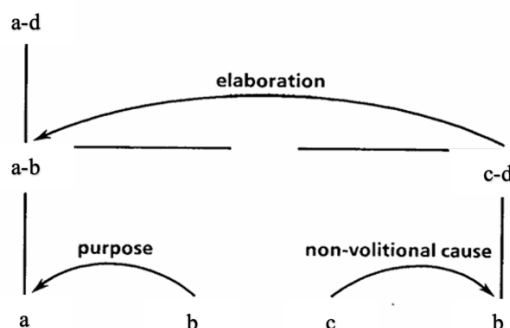
(28)

- a. ...we've been able to mine our own iron ore, coal, manganese, dolomite, all the materials we need.
- b. to make our own steel.
- c. And because we can mine more than we need,
- d. we've had plenty of manganese and iron ore for export.

(Mann & Thompson, 1987, p. 60)

Neste excerto, a frase (28c) exprime a causa não volitiva da situação apresentada em (28d). Como se pode observar, o facto de a empresa conseguir extrair mais recursos do que os necessários conduz, sem envolver qualquer ação intencional específica, à disponibilidade de minério para exportação. As relações entre os diferentes segmentos discursivos são representadas de forma esquemática na Figura 4 apresentada em seguida.

Figura 4 - Relação de causa - Non-volitional cause e purpose



Fonte: (Mann & Thompson, 1985, P. 60)

As causas não volitivas tendem também a incluir os casos em que o núcleo resulta de aplicação de um processo dedutivo, como se pode verificar no exemplo (29) que se segue:

(29)

- a. The elimination of mass poverty is necessary to supply the motivation for fertility control in such countries.

- b. Other countries should assist in this process,
- c. not least because they have a moral obligation to do so.

(Mann & Thompson, 1987, p. 61)

Neste exemplo, (29c) enuncia uma obrigação moral de carácter geral. A partir dessa premissa, deduz-se a obrigação mais específica expressa em (29b), relativa à assistência por parte de outros países. Assim, a relação causal estabelecida não decorre de uma ação volitiva, mas de um raciocínio inferencial que liga uma obrigação geral a uma conclusão prática mais concreta.

Quanto ao grupo de resultado volitivo e resultado não volitivo, deve tomar-se como referência o resultado da relação causal, isto é, aquilo que poderia, ou não, ter sido produzido como consequência de uma determinada situação. Conforme definido por Mann & Thompson (1987), um resultado volitivo ocorre quando uma ação volitiva, isto é, uma ação intencional ou consciente, conduz a uma situação que constitui a consequência direta dessa ação. A relação de resultado volitivo estabelece, assim, uma ligação entre uma ação voluntária e o seu efeito, sendo a situação que origina o efeito deliberadamente realizada pelo agente.

Nesta relação discursiva, o núcleo (N) não apresenta, isoladamente, uma situação claramente caracterizada como ação volitiva. O satélite (S), por sua vez, exprime uma ação volitiva ou uma situação que pode resultar de uma ação volitiva, condicionando a interpretação da combinação N + S. O núcleo apresenta uma situação que pode ser interpretada como causa da situação expressa no satélite, sendo essa situação nuclear mais relevante para os objetivos comunicativos do locutor do que a apresentada no satélite. Nesta configuração, o efeito reconhece que a situação expressa no núcleo pode constituir a causa da ação ou da situação descrita no satélite.

Para ilustrar a relação de resultado volitivo, os autores recorrem a um excerto de um texto que relata a história de uma mulher, Margarida, que afirma que terá de ser operada ao polegar, após concluir que “using thumbs is not the problem, but heredity is, and the end result is no use of thumbs if I don’t do something now” (Mann & Thompson, p. 62). Em consequência desta situação, escrever tornou-se quase

impossível. Por esse motivo, a máquina de escrever foi reparada, com a expectativa de que, após todos esses anos, fosse possível aprender a escrever de forma adequada. A partir da combinação das unidades discursivas presentes no relato, os autores identificam dois grupos que se relacionam entre si, um funcionando como satélite e o outro como núcleo, sendo este último o que exprime o resultado volitivo (cf. Tabela 2).

Tabela 2 - Relação entre argumentos - Núcleo e Satélite

Núcleo	Satélite
a. Using thumbs is not the problem b. but heredity is, c. and the end result is no use of thumbs if I don't do something now.	a. Writing has almost become impossible b. so we had the typewriter serviced c. and I may learn to type decently after all these years.

Fonte: (Mann & Thompson, 1987, pp. 58-62)

As unidades que constituem o satélite (escrever, reparar a máquina de escrever, aprender a escrever de forma adequada), quando articuladas com as unidades do núcleo (a impossibilidade de usar os polegares caso não se intervenha), revelam uma ação voluntária cuja situação causadora do efeito foi deliberadamente realizada pela agente, Margarida.

No que diz respeito ao resultado não volitivo, os autores apresentam a seguinte caracterização. Nesta relação, o satélite (S) exprime uma situação que não resulta de uma ação voluntária, não impondo restrições associadas à intenção do agente. Contudo, na combinação entre núcleo (N) e satélite (S), o núcleo apresenta uma situação que causou a situação descrita no satélite. Deste modo, a situação expressa no núcleo assume um papel mais central para os objetivos comunicativos do locutor do que a apresentada no satélite. Nesta configuração, o efeito reconhece que a situação descrita no núcleo pode ter causado a situação apresentada no satélite.

Com base nesta definição, o resultado não volitivo tende a associar-se a acontecimentos que ocorrem independentemente da vontade do agente, sendo consequências das suas ações ou de determinadas circunstâncias, mas não como resultados intencionalmente realizados ou desejados. Trata-se, assim, de efeitos que se produzem sem uma intenção deliberada por parte de quem os desencadeia.

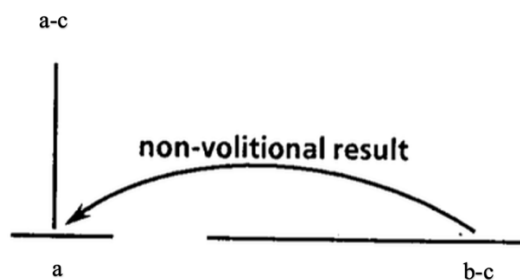
Esta relação pode ser observada no excerto seguinte (cf. 30), retirado de uma notícia da BBC World Service, bem como na Figura 5 que se apresenta de seguida:

(30)

- a. The blast, the worst industrial accident in Mexico's history, destroyed the plant and most of the surrounding suburbs.
- b. Several thousand people were injured,
- c. and about 300 are still in the hospital.

(Mann & Thompson, 1987, p. 62)

Figura 5 - Relação de causa - Non-Volitional result



Fonte: (Mann & Thompson, 1985, P.63)

Neste caso, a unidade a) identifica-se como núcleo, uma vez que o desastre resultou na destruição da fábrica e dos subúrbios circundantes, enquanto as unidades b) e c) apresentam os resultados não volitivos do acidente. O resultado não volitivo consiste no facto de várias pessoas terem ficado feridas e de muitas terem sido hospitalizadas, sem que tal tivesse sido intencional ou desejado. Trata-se, assim, de consequências involuntárias decorrentes da explosão.

A última relação pertencente à classe das relações de causa na teoria de Mann & Thompson (1987) é a finalidade (*Purpose*). Esta relação estabelece uma ligação entre duas situações, na qual a situação nuclear é realizada com o propósito de provocar ou permitir que a situação apresentada no satélite ocorra. De acordo com a definição proposta, o núcleo (N) exprime uma atividade, enquanto o satélite (S) apresenta uma

situação que ainda não foi realizada. Contudo, quando considerados em conjunto, na combinação N + S, o satélite passa a representar uma situação a ser concretizada por meio da atividade expressa no núcleo. O efeito ou resultado reconhece, assim, que a atividade apresentada em N é iniciada com o objetivo de realizar a situação expressa em S.

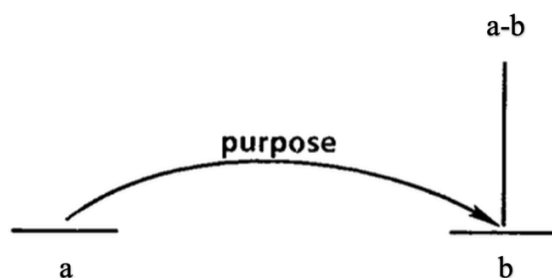
Para uma melhor compreensão desta relação, considere-se o excerto final de um anúncio de disquetes, apresentado em (31).

- (31) a. To see which Syncom diskette will replace the ones you're using now,
b. send for our free "Flexi-Finder" selection guide and the name of the supplier nearest you.

(Man & Thompson, 1985, P. 64)

Neste exemplo, a situação nuclear corresponde à atividade de solicitar o guia gratuito, que é realizada com o propósito de concretizar a situação apresentada no satélite, isto é, permitir ao leitor descobrir qual a disquete que substituirá a que utiliza atualmente. A ação enunciada transmite, assim, de forma explícita, o objetivo que orienta a realização da atividade descrita. A representação desta análise encontra-se ilustrada na Figura 6.

Figura 6 - Relação de causa - Purpose



Fonte: (Mann & Thompson, 1985, P. 64)

A relação de finalidade é considerada neutra no que diz respeito à volição, o que significa que pode associar-se tanto a ações intencionais quanto a situações que não envolvem uma vontade consciente ou planeada..

2.3. *Segmented Discourse Representation Theory* (Asher & Lascarides, 2003)

Para além dos autores já abordados nos subcapítulos anteriores deste capítulo, importa destacar o contributo de Asher e Lascarides (2003) com a *Segmented Discourse Representation Theory* (doravante SDRT). Estes autores desenvolvem uma teoria da interpretação do discurso que estende a semântica dinâmica, integrando nela as relações retóricas como parte das formas lógicas do discurso. A SDRT atribui um papel central às relações retóricas no processo de interpretação e postula que o significado de um discurso depende da sua estrutura retórica e interage com ela.

Nesta perspetiva, o discurso é entendido como uma estrutura composta por relações discursivas que ligam proposições ou conteúdos enunciativos, formando uma rede de significados interligados. Um discurso é considerado coerente na medida em que as proposições que o constituem se encontram relacionadas entre si de forma significativa, ou seja, “the rhetorical role of a bit of information depends on how it relates to some other bit of information in the context” (Asher & Lascarides, 2003, p. 3), condição necessária para a manutenção da coerência discursiva. Um discurso torna-se incoerente quando uma proposição introduzida não se encontra ligada de forma significativa a qualquer outro elemento do discurso (Asher & Lascarides, 2003).

Contudo, como muitos aspetos do significado do discurso não são explicitamente codificados, os falantes têm de recorrer a inferências baseadas no conhecimento linguístico (como a sintaxe e a semântica composicional) e no conhecimento do mundo. Como sublinham Asher & Lascarides,

...people don't make explicit many aspects of what we take to be part of the discourse's meaning. In particular, they don't always make explicit the information about rhetorical function or the implicatures of rhetorical function. Instead, interpreters infer, and are expected to infer, information over and above what's derivable from the meaning of the individual words and syntax. While this happens even when

constructing the meaning of clause, it is pervasive when constructing the meaning of a multi-sentence discourse (Asher & Lascarides 2003, p. 1).

Deste modo, diversos fatores influenciam positiva ou negativamente a compreensão e a interpretação de um discurso, entre os quais se incluem fenómenos textuais como a anáfora individual, a anáfora de entidades abstratas, a elipse verbal, a estrutura temporal das eventualidades descritas, as inferências de ligação (*bridging inferences*), as pressuposições, a ambiguidade lexical e as inferências subjacentes ao diálogo, nomeadamente as implicaturas conversacionais (Asher & Lascarides, 2003).

Segundo estes autores, o discurso apresenta uma estrutura hierárquica, refletindo a possibilidade de os falantes variarem a granularidade com que descrevem situações e acontecimentos. Esta estrutura discursiva condiciona também a organização temporal dos eventos, contribuindo para a interpretação global do discurso. O sentido de um texto resulta, assim, da interação entre a sua estrutura e os significados das unidades que o compõem, sendo os diferentes enunciados ou orações articulados de modo a produzir coerência. Essas articulações podem ser explicitamente marcadas por conectores discursivos, conjunções, preposições ou outros elementos linguísticos, mas podem igualmente ser inferidas.

É com base nestes pressupostos que Asher & Lascarides (2003) distinguem cinco grandes classes de relações retóricas:

- relações de conteúdo, que ocorrem em orações declarativas, interrogativas e imperativas;
- relações da estrutura textual, associadas à forma como as situações são organizadas e representadas no discurso;
- relações de nível cognitivo, que correspondem a relações inferidas no diálogo, geralmente deduzidas a partir dos atos ilocutórios, das intenções e das crenças dos interlocutores;
- relações divergentes, que envolvem contrastes ou incompatibilidades na estrutura dos constituintes discursivos;
- relações metaconversacionais, que se estabelecem entre o conteúdo dos enunciados e os objetivos dos atos de fala.

O estudo de Asher & Lascarides (2003) identifica um conjunto alargado de relações retóricas, entre as quais se incluem Narração, Continuação, Enquadramento, Resultado, Alternância, Explicação e Elaboração, entre outras. Tendo em conta o foco desta investigação, serão desenvolvidas de forma mais detalhada as relações associadas à causalidade, em particular as relações de Explicação e de Resultado.

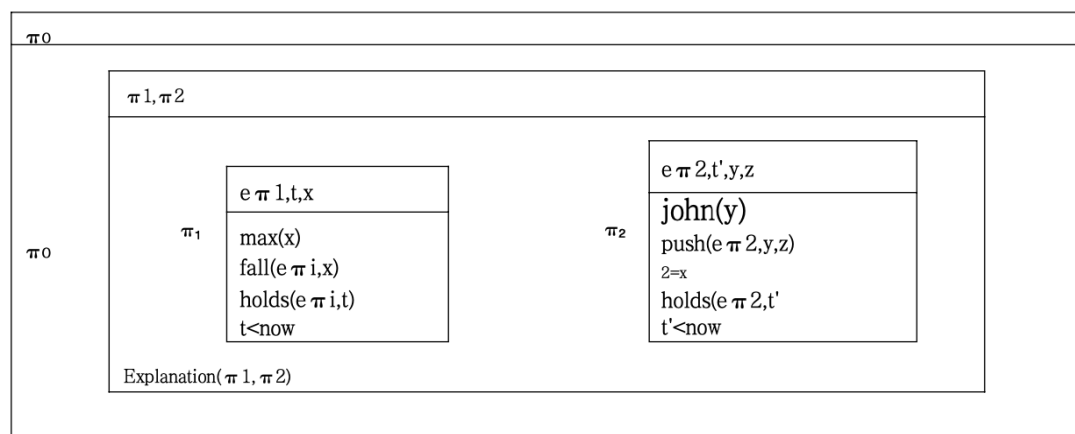
Quanto à relação retórica de Explicação, os enunciados ou argumentos acrescentados à oração principal fornecem informação justificativa que legitima essa oração, complementando o seu sentido. Neste caso, a proposição β é logicamente dependente da proposição α , estabelecendo-se entre ambas uma relação causal entre os eventos descritos. Observe-se o exemplo seguinte:

(32) Max fell. John pushed him.

Asher & Lascarides (2003, p. 6)

Neste exemplo, a proposição de que John empurrou Max explica por que razão a proposição de que Max caiu é verdadeira. Além disso, o evento de empurrar ocorre temporalmente antes do evento de cair, o que reforça a interpretação causal da relação. A Figura 7 contém a *Segmented Discourse Representation Structure* (SDRS) que formaliza a interpretação desta frase.

Figura 7 - Representação das frases em estruturas de representação discursiva



Fonte: Silvano (2010, p.210)

No que diz respeito às relações retóricas de Resultado, de acordo com Asher & Lascarides (2003), esta relação discursiva estabelece uma ligação de resultado entre dois eventos, em que α corresponde à causa e β ao efeito. A relação de Resultado constitui uma estrutura semântica que liga uma causa a um efeito, exigindo que a causa preceda temporalmente o efeito e que ambos os eventos sejam considerados verdadeiros no discurso. Asher & Lascarides (2003) caracterizam esta relação como “dual” da relação de Explicação. Como afirmam os autores:

In words, Result stipulates that if β is to be attached to α and the content of this attachment is to be assigned the label λ , and there's evidence in the discourse σ that α caused β , then normally Result (α , β) holds (at λ) regardless of the aspectual classes of α and β (i.e. whether the main eventualities are classified as events or states) (Asher & Lascarides, 2003, p. 206).

Deste modo, entende-se que uma relação de Resultado, tal como apresentada num discurso σ , é representada pela proposição de que β constitui um resultado da proposição α . Em regra, existe uma ligação causal identificável entre estas duas proposições, sendo irrelevante, para a definição da relação, a classe aspectual dos conteúdos envolvidos, quer se trate de eventos quer de estados. Em termos temporais, o efeito estabelece uma relação de posterioridade com a causa.

2.4. Extensão da SDRT (Cunha & Silvano, 2009; Silvano, 2010)

A distinção entre relações retóricas intrínsecas e relações retóricas extrínsecas, proposta por Cunha & Silvano (2009), insere-se no quadro teórico da *Segmented Discourse Representation Theory*, de Asher & Lascarides (2003), constituindo simultaneamente um desenvolvimento dessa abordagem. Em Asher & Lascarides (2003), as relações retóricas são concebidas como ligações semântico-pragmáticas entre proposições, responsáveis pela coerência discursiva, sendo organizadas segundo diferentes níveis, como o nível do conteúdo e o nível da estrutura do discurso. No entanto, embora a SDRT reconheça a diversidade funcional das relações retóricas, não estabelece uma distinção explícita baseada no grau de interdependência semântica entre as situações envolvidas.

É neste ponto que a proposta de Cunha & Silvano (2009) introduz um contributo relevante. Os autores distinguem entre relações retóricas extrínsecas e intrínsecas com

base no grau de ligação semântica entre as eventualidades, permitindo explicar a variedade de leituras temporais e discursivas observadas, em particular, em frases complexas introduzidas pela conjunção *quando* em português europeu. Esta distinção assenta na ideia de que nem todas as relações discursivas são inferidas a partir do mesmo tipo de informação nem com o mesmo grau de dependência semântica.

As relações extrínsecas caracterizam-se por uma fraca interdependência semântica entre as situações envolvidas. Nestes casos, a ligação discursiva estabelece-se por defeito ao nível da organização externa das predicações, sendo sobretudo determinada pela ordem temporal dos eventos ou pela sua simples coexistência no discurso. A inferência destas relações depende maioritariamente de fatores estruturais e temporais, e não de elos semânticos fortes entre os conteúdos proposicionais. Exemplos típicos de relações extrínsecas são as relações de Narração e de Enquadramento, em que os eventos se sucedem ou se sobrepõem temporalmente, sem que uma situação constitua causa, resultado ou parte constitutiva da outra.

Em contraste, as relações intrínsecas envolvem uma forte interdependência semântica entre as situações. Nestas relações, a conexão discursiva resulta de elos semânticos profundos estabelecidos ao nível da estrutura interna das predicações e não apenas da sua ordenação temporal. Relações como Explicação, Resultado e Elaboração pertencem a este grupo. A inferência destas relações depende da interação de várias fontes de informação, nomeadamente a semântica composicional, o léxico e o conhecimento do mundo, sendo a ordenação temporal das eventualidades uma consequência da relação retórica inferida, e não o seu ponto de partida.

Para ilustrar de modo mais concreto as diferenças entre duas das relações intrínsecas de causa e efeito, Cunha & Silvano (2009) comparam as relações de Explicação e Resultado. Embora ambas envolvam uma noção de causalidade entre as situações, elas diferem na estrutura discursiva: na Explicação, a oração subordinada introduzida por *quando* fornece a causa para o evento da oração principal, enquanto no Resultado a subordinada expressa o efeito decorrente do evento principal. A Tabela 4 resume essa distinção com definições e exemplos.

Tabela 3 - Relação de Explicação e de Resultado

Relações retóricas	Definição semântica	Exemplos	Consequências temporais
Explicação	A segunda situação descreve a causa subjacente à ocorrência da primeira situação.	Quando o João caiu, escorregou numa casca de banana.	e1:caiu e2: escorregou e1<e2
Resultado	A segunda situação descreve o efeito da ocorrência da primeira situação.	Quando o João escorregou numa casca de banana, caiu.	e1: escorregou e2: caiu e2< e1

Fonte: Cunha & Silvano (2009, p. 240)

Nos exemplos acima, observa-se a diferença na ordenação temporal conforme a relação retórica: na explicação, a ação de *escorregar* (subordinada) ocorreu antes e levou à ação de *cair* (principal), descrevendo um cenário em que a segunda oração fornece a explicação causal para o evento da primeira, logo, trata-se de uma Explicação. na resultado, a queda narrada na oração principal acontece após o escorregar mencionado na subordinada (ou seja, o escorregar foi imediatamente seguido da queda); neste caso, a subordinada relata um acontecimento resultante do evento principal, estabelecendo, portanto, uma relação de Resultado. Este par ilustrativo evidencia como a mesma sequência de eventos (*João escorrega* e *João cai*) pode ser descrita de formas diferentes com *quando*, originando interpretações distintas conforme a estrutura discursiva: ou enfatizamos o efeito primeiro e depois revelamos a causa (Explicação), ou apresentamos a causa inicial seguida do efeito consequente (Resultado).

A proposta de Cunha & Silvano (2009) é compatível com a SDRT, na medida em que mantém a concepção das relações retóricas como relações inferidas a partir de múltiplas fontes de informação, tal como defendido por Asher & Lascarides (2003). Contudo, ao introduzir a distinção entre relações intrínsecas e extrínsecas, os autores clarificam o papel relativo dessas fontes no processo inferencial: nas relações extrínsecas, a inferência é largamente determinada pela organização temporal e estrutural do discurso e nas relações intrínsecas, a inferência assenta sobretudo em elos semânticos fortes, sendo a temporalidade subordinada à relação discursiva subjacente.

Silvano (2010) baseia-se igualmente na proposta de Asher e Lascarides (2003) para a análise de frases complexas, incidindo, em particular, sobre orações adverbiais temporais, causais, concessivas, condicionais e orações finais. O objetivo central da autora consiste em compreender de que modo estas orações se organizam ao nível temporal e como se articulam no plano das relações discursivas.

No decurso da sua análise, Silvano (2010) identifica alguns problemas na aplicação direta da *Segmented Discourse Representation Theory*, de Asher & Lascarides (2003), à subordinação adverbial, ainda que reconheça tratar-se de uma das teorias mais sólidas no tratamento das relações retóricas. A este propósito, a autora sublinha que a teoria carece de desenvolvimentos adicionais que lhe permitam dar conta da diversidade e da especificidade dos dados empíricos analisados, salientando a necessidade de uma caracterização mais fina das propriedades das relações discursivas, de modo a alargar a abrangência explicativa do enquadramento teórico.

Um dos problemas observados prende-se com a posição das orações subordinadas, quer em posição inicial quer em posição final, e com as consequências dessa posição para a direccionalidade da ligação entre os argumentos das relações discursivas. A SDRT prevê a restrição da Fronteira da Direita, segundo a qual a ligação retórica deve seguir o percurso que vai do último rótulo introduzido até ao constituinte mais alto da estrutura discursiva. Este procedimento aplica-se de forma relativamente clara a exemplos constituídos por frases simples, mas levanta dificuldades quando se trata de frases complexas (Silvano, 2010). Coloca-se, assim, a questão de saber se a direccionalidade das relações discursivas se mantém inalterada nas frases complexas ou se varia em função do tipo de oração adverbial envolvida. Após análise dos dados, a autora defende que a direccionalidade das relações causais, concessivas, condicionais e finais deve ser estabelecida da oração subordinada para a oração principal, independentemente da ordem linear, enquanto, nas orações temporais, a direccionalidade deve ser da oração principal para a subordinada. O tratamento diferente das orações subordinadas adverbiais no que diz respeito à direccionalidade justifica-se pelo papel determinante desempenhado pelos conectores causais,

concessivos, condicionais e finais na determinação da relação retórica, por um lado, e por outro lado, pela função de localização temporal assumida pelas orações temporais.

Um outro problema observado por Silvano (2010) refere-se à ordem dos argumentos e ao impacto da posição da oração subordinada na interpretação discursiva. A autora argumenta que a colocação da oração adverbial em posição inicial ou final não é irrelevante para a organização estrutural do discurso. Mesmo quando se infere a mesma relação retórica ao nível do conteúdo, essas posições acrescentam diferentes valores ao significado da frase. Torna-se, por isso, necessário considerar não apenas o nível do conteúdo, mas também o nível estrutural na determinação das relações retóricas. Embora Asher & Lascarides (2003) distingam relações ao nível estrutural, Silvano (2010) mostra que as relações propostas não representam de forma adequada e completa o significado das orações adverbiais em posição inicial e final. Por isso, propõe mais duas relações discursivas: *Frame*, que é inferida quando a oração adverbial ocupa uma posição inicial, estabelecendo um enquadramento para a situação da oração principal; e *Specification*, que é inferida quando a oração adverbial ocupa uma posição final e fornece informações adicionais (motivo, finalidade,...) sobre a situação descrita pela oração principal.

No que diz respeito às relações de Explicação e Resultado, o foco deste trabalho, Silvano (2010) analisa-as, tal como Cunha & Silvano (2009), como relações intrínsecas de causalidade entre situações descritas em enunciados distintos. Em termos formais, Silvano (2010) define Explicação (α , β) como a relação em que uma das proposições expressa uma situação que é a causa ou razão da situação representada pela outra proposição. Já Resultado (α , β) é definido simetricamente como ocorrendo quando uma das proposições expressa uma situação que é o efeito ou consequência da situação representada pela outra. Estas relações são mutuamente exclusivas dada sua natureza causal semelhante.

Silvano (2010) mostra que as orações subordinadas adverbiais causais em português tipicamente estabelecem uma relação discursiva de Explicação. As orações causais introduzidas por conetores como *porque*, *visto que*, *dado que*, *já que* ou *como* (em sentido causal), entre outros, apresentam uma proposição que justifica ou explica

a situação da oração principal. A autora acrescenta que mesmo quando a oração causal antecede a principal (estrutura inversa), os conectores causais sinalizam inequivocamente uma relação explicativa, não devendo a oração ser reinterpretada como Resultado apenas com base na ordem linear.

Para além das causais, Silvano (2010) demonstra que algumas orações temporais podem também veicular relações de Explicação, caso a interpretação ultrapasse a mera sequência temporal e incorpore um nexos causal implícito. Um exemplo ilustrativo é o seguinte retirado de Silvano (2010, p. 287):

(33) O Rui tinha comido três hambúrgueres quando se sentiu mal. (Explicação)

Quanto à relação de Resultado, Silvano (2010) mostra que ela pode ser inferida quando ocorrem orações adverbiais finais introduzidas por conectores como *para (que)* ou *a fim de (que)*. Para além destas, também as orações temporais podem implicar uma relação de Resultado quando a sequência descrita sugere causalidade direta. Novamente, a conjunção *quando* pode ser interpretada dessa forma em contextos apropriados. Por exemplo, Silvano (2010, p. 287) analisa a frase:

(34) O João caiu quando tropeçou no cabo da televisão. (Resultado)

Nesta frase, a oração subordinada temporal (*quando tropeçou no cabo da televisão*) descreve um acontecimento cuja ocorrência provoca o evento da oração principal (João cair). A ligação inferida é de causa → consequência, ou seja, de Resultado, dado que tropeçar levou efetivamente à queda.

Em suma, das cinco classes de subordinadas tradicionais (temporais, causais, concessivas, condicionais e finais), Silvano (2010) conclui que apenas as orações causais e algumas temporais estabelecem relações de Explicação, e apenas as orações finais e algumas temporais estabelecem relações de Resultado.

2.5. Penn Discourse Treebank (Prasad et al., 2008; 2010; 2019)

A Universidade da Pensilvânia tem vindo a desenvolver o projeto *Penn Discourse Treebank* (PDTB), cujo objetivo consiste na recolha e análise de relações discursivas a partir de um corpus de textos em inglês anotado com etiquetas de parte do discurso e outras informações linguísticas. Segundo Prasad et al. (2019), o projeto PDTB assenta na ideia de que as relações discursivas se fundamentam num conjunto identificável de palavras ou expressões explícitas, os chamados conectores discursivos, ou, em alternativa, na simples adjacência de duas frases. O PDTB tem sido amplamente utilizado por investigadores da comunidade de Linguística e Processamento de Linguagem Natural. Além disso, tem estimulado o desenvolvimento de recursos semelhantes noutras línguas e domínios. O facto de não adotar uma perspetiva teórica específica permite um uso vasto deste recurso, como referem Prasad et al. (2008):

The PDTB is also unique in adopting a theory-neutral approach to the annotation, making no commitments to what kinds of high-level structures may be created from the lowlevel annotations of relations and their arguments. This approach has the appeal of allowing the corpus to be useful for researchers working within different frameworks while at the same time providing a resource to validate the various existing theories of discourse structure (Mann and Thompson, 1988; Webber and Joshi, 1998; Wolf and Gibson, 2005). This theory neutrality also permits investigation of the general question of how structure at the sentence level relates to structure at the discourse level, at least that part of the discourse structure that is “parallel” to the sentence structure (Lee et al., 2006) (Prasad et al., 2008, p. 2961).

O *Penn Discourse Treebank* 3.0 corresponde à terceira versão do projeto, cujo objetivo foi anotar secções dos artigos do *Wall Street Journal* (WSJ) do Treebank-2 (LDC95T7) com relações discursivas. A tarefa inicial consistiu na anotação de cada palavra com a respetiva etiqueta de parte do discurso. Na versão 3.0, que constitui o foco desta revisão, foram anotados mais de 13 000 tokens adicionais, foram normalizadas determinadas anotações por pares, foram introduzidos novos sentidos e o corpus foi submetido a uma série de verificações de coerência.

O PDTB foi o primeiro projeto a adotar uma abordagem lexicalmente fundamentada para a anotação das relações discursivas. Quando as relações discursivas são realizadas explicitamente no texto, são anotadas através da identificação dos itens lexicais necessários, conectores discursivos, o que facilita a sua identificação automática.

Outro aspeto relevante salientado por estes autores prende-se com o facto de o PDTB fornecer etiquetas de sentido para cada relação discursiva, seguindo um esquema hierárquico de unidades de sentido presentes no texto. A anotação dos sentidos evidencia a polissemia dos conectores discursivos, aspeto particularmente relevante para tarefas de desambiguação semântica, o que torna o PDTB um recurso valioso para o estudo das relações discursivas.

As relações discursivas “are assumed to have two and only two arguments, called Arg1 and Arg2. By convention, Arg2 is the argument syntactically associated with the relation, while Arg1 is the other argument” (Prasad et al., 2010, p. 1024). Especificam também que “In the case of Explicit connectives, Arg2 is the argument to which the connective is syntactically bound, and Arg1 is the other argument. In the case of relations between adjacent sentences, Arg1 and Arg2 reflect the linear order of the arguments, with Arg1 before Arg2” (Prasad et al., 2008, p. 2962).

Sendo o foco deste trabalho as relações discursivas de causalidade, centrar-nos-emos neste tipo de relações tal como são definidas no âmbito do PDTB. No que diz respeito às relações explícitas de causalidade, estas são normalmente introduzidas por conectores discursivos, como *since* ou *because*, que estabelecem de forma direta a relação entre dois segmentos textuais. Os exemplos seguintes ilustram este tipo de relação explícita:

(35) It was a far safer deal for lenders since NWA had a healthier cash flow and more collateral on hand.

(36) Domestic car sales have plunged 19% since the Big Three ended many of their programs Sept. 30.

(Prasad et al., 2008, p. 2964)

Contudo, um aspeto central da abordagem do PDTB é o reconhecimento de que as relações discursivas nem sempre são expressas por meio de conectores explícitos. Em muitos casos, essas relações são inferidas pelo leitor com base no conteúdo semântico e no encadeamento discursivo dos segmentos envolvidos. Segundo Prasad et al. (2019),

estas relações são designadas como implícitas. Considere-se o exemplo seguinte de Prasad et al. (2008, p. 2963):

(37) But a few funds have taken other defensive steps. Some have raised their cash positions to record levels. Implicit = BECAUSE High cash positions help buffer a fund when the market falls.

Neste caso, estabelece-se uma relação causal entre o aumento das posições de tesouraria e a função protetora dessas posições em contextos de queda do mercado. Embora não exista no texto nenhum conector causal explícito, a relação de causalidade é naturalmente inferida. No processo de anotação do PDTB, esta inferência é representada pela introdução hipotética de um conector discursivo adequado, razão pela qual se fala em conetivo implícito.

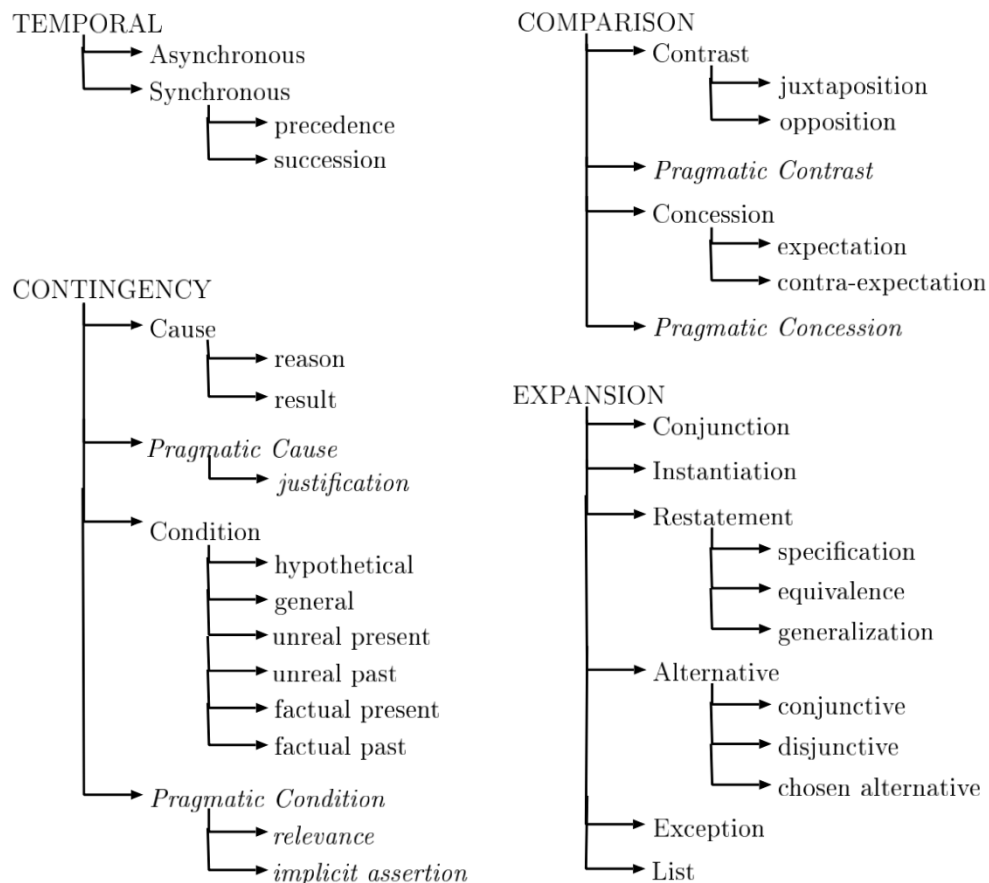
Para além das relações explícitas e implícitas, o PDTB identifica ainda os casos de lexicalização alternativa (Alternative Lexicalization, ou AltLex). Trata-se de situações em que a relação discursiva é expressa por uma construção lexical específica que desempenha a função de um conector, tornando desnecessária, ou mesmo redundante, a inserção de um conector canónico. Observe-se o exemplo seguinte de Prasad et al. (2010, p. 1025):

(38) Now, GM appears to be stepping up the pace of its factory consolidation to get in shape for the 1990s. One reason is mounting competition from new Japanese car plants in the U.S. that are pouring out more than one million vehicles a year at costs lower than GM can match. (Contingency: Cause: reason).

Neste caso, a expressão “One reason is” funciona como marcador da relação causal entre os dois segmentos discursivos. A introdução de um conector como *because* resultaria numa redundância informacional, uma vez que a relação de causa já se encontra lexicalmente codificada. Por esse motivo, esta construção é anotada no PDTB como AltLex.

Para além da identificação dos argumentos discursivos (Arg1 e Arg2) e dos diferentes tipos de marcadores explícitos, implícitos e AltLex, o PDTB procede também à anotação dos sentidos ou relações semânticas associadas a essas ligações discursivas. O sistema de sentidos encontra-se organizado de forma hierárquica, em três níveis: classe, tipo e subtipo. O nível mais alto, o das classes, compreende quatro grandes categorias semânticas: Temporal, Contingency, Comparison e Expansion (Prasad et al., 2008, p. 2964), como se ilustra na figura (8).

Figura 8 - Hierarquia de relações discursivas



Fonte: (Prasad et al, 2008, p. 2965)

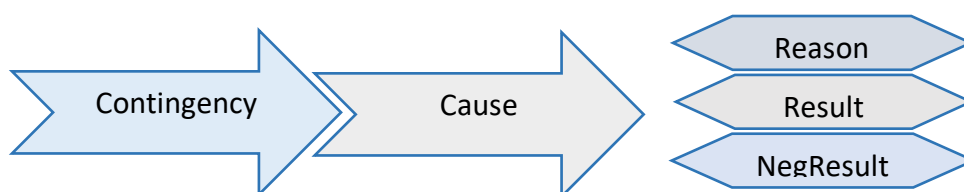
De acordo com Prasad et al. (2008), a classe Contingency, ilustrada na figura anterior, é utilizada quando a situação descrita por um dos argumentos fornece a razão, explicação ou justificação para a situação descrita pelo outro argumento. No interior desta classe, o tipo Cause aplica-se aos casos em que as situações descritas em Arg1 e

Arg2 se encontram ligadas por uma relação causal, sem que essa relação assuma a forma de uma condição, isto é, sem configurar uma relação condicional.

Na versão 2 do PDTB, a relação de causa é subdividida em dois subtipos. O primeiro, *reason*, é utilizado quando o conector indica que a situação expressa em Arg2 é interpretada como a causa da situação descrita em Arg1, como acontece tipicamente com conectores como *because* (Prasad et al., 2008). O segundo subtipo, *result*, aplica-se quando o conector indica que a situação expressa em Arg2 corresponde ao resultado da situação apresentada em Arg1, como ocorre com expressões do tipo *as a result* (Prasad et al., 2008). Na versão 3 do PDTB, foi introduzido um novo subtipo de sentido sob o tipo Cause, designado NegResult. Este subtipo é utilizado quando Arg1 fornece a razão, explicação ou justificação que impede a concretização do efeito mencionado em Arg2. Segundo Prasad et al. (2019, p. 20), este sentido foi introduzido especificamente para dar conta de construções léxico-sintáticas do tipo *too X to Y* (*The weather is too bad to go out*), nas quais o primeiro segmento exprime uma condição que inviabiliza o resultado potencial expresso no segundo. Esta reformulação teve como objetivo aprofundar a análise das relações intrassentenciais (Intra-S), isto é, das relações discursivas que ocorrem no interior de uma única frase.

Os subtipos de Cause atualmente reconhecidos no PDTB-3 encontram-se organizados na figura (9).

Figura 9 - Os subtipos de Causa na classe Contingency



Fonte: adaptado de (Prasad et al., 2019 p. 17)

Os exemplos seguintes ilustram cada um dos três subtipos referidos:

(39) **Reason:**

But service on the line is expected to resume by noon today. (Implicit=since)
“We had no serious damage on the railroad,” said a Southern Pacific
spokesman. [wsj 1803]

(Prasad et al., 2019, p. 19)

(40) Result:

The bill would then declare that the debt is equity and therefore isn’t
deductible. [wsj 1822]

(Prasad et al., 2019, p. 20)

(41) NegResult:

Friday’s stock market sell-off came too late for many investors to act. [wsj
2306]

(Prasad et al., 2019, p. 20)

Nestes exemplos, observa-se, respetivamente, uma relação em que a causa é apresentada como justificação do evento descrito, uma relação em que o segundo segmento exprime o resultado do primeiro e, por fim, uma relação em que o evento inicial impede a ocorrência do efeito potencialmente esperado.

2.6. Core-Discourse Relations (ISO 24617-8, 2016)

A *International Organization for Standardization* (ISO) é uma associação mundial de organismos nacionais de normalização, cujo principal objetivo consiste na elaboração de normas internacionais através do trabalho desenvolvido em comités técnicos especializados. Cada organismo membro interessado numa determinada área temática pode propor a criação de um comité técnico, no qual participam especialistas designados pelas entidades associadas. Para além dos organismos nacionais, organizações internacionais, governamentais e não governamentais, em articulação com a ISO, podem igualmente participar nos trabalhos de normalização. O processo de normalização assenta na análise crítica das diferentes abordagens teóricas existentes sobre um determinado domínio, a partir das quais são definidos padrões comuns, formalizados em normas internacionais.

No domínio dos recursos linguísticos, a norma ISO 24617, intitulada *Language Resource Management – Semantic Annotation Framework* (SemAF), define um quadro de referência para a anotação semântica, organizado em vários níveis, incluindo, entre outros, os níveis temporal, referencial, espacial e funcional. Este enquadramento permite a anotação semântica de diferentes tipos de entidades e relações linguísticas, promovendo a interoperabilidade entre recursos e sistemas. Como salientado por Tomaszewska et al. (2024), esta abordagem modular facilita a integração de diferentes perspetivas teóricas num quadro comum de anotação.

A norma ISO 24617-8, publicada em 2016, propõe uma abordagem interoperável para a anotação de relações discursivas ao nível local. Esta norma sistematiza princípios gerais que emergem de diferentes quadros teóricos, evidenciando fortes convergências no que diz respeito à descrição semântica das relações discursivas. Em vez de impor um conjunto fechado de relações, a ISO 24617-8 propõe um inventário aberto e extensível de relações fundamentais, compatibilizando e mapeando definições provenientes de múltiplas tradições teóricas.

Para esse efeito, a ISO mobiliza contributos de várias teorias de referência, nomeadamente a *Rhetorical Structure Theory* (RST), a *Segmented Discourse Representation Theory* (SDRT), o *Penn Discourse Treebank* (PDTB), a *Hobbs' Theory of Discourse Coherence* (HTDC) e a *Cognitive Approach to Coherence Relations* (CCR), com o objetivo de estabelecer normas e padrões reutilizáveis por investigadores que pretendam dedicar-se ao estudo de relações discursivas ou criar recursos linguísticos. A ISO 24617-8 define um conjunto de conceitos fundamentais para a compreensão global das teorias das relações discursivas, entre os quais se destacam:

- discourse - “sequence of clauses or sentences in written text or of utterances in oral speech”;
- situation - “eventuality, fact, proposition, condition, belief or dialogue act, that can be realized by a linguistically simple or complex expression, “such as a clause, a nominalization, a sentence/utterance, or a discourse segment consisting of multiple sentences or utterances”;
- discourse relation – “relation between two situations mentioned in a discourse”;
- discourse connective – “word or multi-word expression expressing a discourse relation”;
- low-level discourse structure – “representation of discourse structure that only specifies local dependencies between a discourse relation and its arguments, without further specifying any links or dependencies across these local structures”.

Neste enquadramento, a norma ISO 24617-8 define as relações discursivas, também designadas relações de coerência ou relações retóricas, como relações, explícitas ou implícitas, entre situações mencionadas num discurso. Estas relações são consideradas essenciais para uma compreensão completa do discurso, indo além do significado veiculado por frases ou orações isoladas. A estrutura discursiva e as relações discursivas são igualmente apontadas como componentes centrais para tarefas de processamento de linguagem natural, tais como sumarização, resposta a perguntas complexas, geração automática de texto, tradução automática, análise de opinião e recuperação de informação (ISO 24617-8, 2016).

Esta definição evidencia dois aspetos fundamentais: por um lado, a natureza semântica das relações discursivas; por outro, a necessidade de uma representação estruturada do discurso. Quanto ao significado discursivo, a ISO sublinha que, num discurso coerente escrito ou oral, as situações mencionadas (eventos, estados, factos, proposições ou atos de diálogo) se encontram semanticamente ligadas por relações causais, temporais, contrastivas, entre outras. Estas relações podem ocorrer entre unidades de diferentes dimensões, desde nominalizações e orações até parágrafos ou segmentos dialogais, e podem também envolver situações não explicitamente expressas, mas inferíveis a partir do contexto e do conhecimento do mundo (ISO 24617-8, 2016).

A análise e interpretação de um discurso envolvem, portanto, processos complexos, que articulam a organização sequencial das ideias com a inferência das relações que as ligam. O estudo das relações discursivas implica, por isso, a consideração de múltiplos aspetos adicionais, para além da descrição semântica e da representação estrutural, nomeadamente: variantes pragmáticas das relações, hierarquias semânticas, múltiplas inferências entre segmentos, assimetria relacional, importância relativa dos argumentos, aridade, forma sintática e adjacência dos argumentos, desencadeadores de relações, atribuição discursiva, relações baseadas em entidades, ausência de relação discursiva, bem como os pressupostos e extensões do núcleo DR-core e o respetivo metamodelo.

Relativamente à representação da estrutura do discurso, a norma reconhece a

coexistência de diferentes modelos formais. Alguns quadros teóricos, como a RST, concebem o discurso como uma estrutura em árvore que abrange a totalidade do texto. Outros, como o *Discourse GraphBank*, baseado na HTDC, admitem grafos gerais com múltiplos pares e cruzamentos. Já os corpora DISCOR e ANNODIS, fundamentados na SDRT, recorrem a grafos acíclicos dirigidos, que permitem múltiplos pares, mas excluem cruzamentos (ISO 24617-8, 2016). Para além destas abordagens, a ISO identifica representações consideradas pré-teóricas ou teoricamente neutras, como o PDTB e os modelos baseados na CCR, nas quais as relações discursivas são anotadas individualmente, juntamente com os seus argumentos, sem serem integradas numa estrutura global que represente o texto como um todo (ISO 24617-8, 2016).

Dado que estas diferentes representações estruturais são difíceis de conciliar num único modelo, a ISO 24617-8 adota uma perspectiva pré-teórica assente na anotação de baixo nível das relações discursivas. Parte-se do princípio de que as relações locais podem ser anotadas de forma mais fiável, podendo posteriormente ser integradas em estruturas de nível superior, árvores ou grafos, de acordo com o quadro teórico escolhido pelo investigador (ISO 24617-8, 2016).

Quando se pretende desenvolver um sistema unificado que seja compatível com diferentes enquadramentos teóricos, torna-se, assim, fundamental considerar tanto as teorias consolidadas como as abordagens pré-teóricas e neutras, assegurando a interoperabilidade do sistema.

A ISO 24617-8 (2016) propõe vinte tipos de relações discursivas (cf. Tabela 4), definidos a partir da comparação e compatibilização de cinco taxonomias semânticas amplamente reconhecidas: o *Penn Discourse Treebank*, a *Hobbs' Theory of Discourse Coherence*, a *Rhetorical Structure Theory*, a *Segmented Discourse Representation Theory* e a *Cognitive Approach to Coherence Relations*. Esta convergência constitui um dos principais contributos da norma para a investigação linguística e para o desenvolvimento de recursos interoperáveis em processamento de linguagem natural.

Tabela 4 - Relações discursivas propostas pela ISO 24617-8 (2016)

	ISO 24617-8	Symmetry	Relations and argument role definitions
1.	Cause	Asymmetric	Arg2 is an explanation for Arg1.
2.	Condition	Asymmetric	Arg2 is an unrealized situation which, when realized, would lead to Arg1.
3.	Negative condition	Asymmetric	Arg2 is an unrealized situation which, when “not” realized, would lead to Arg1.
4.	Purpose	Asymmetric	Arg2 is the goal or purpose of the situation described by Arg1.
5.	Manner	Asymmetric	Arg2 describes how Arg1 comes about or occurs.
6.	Concession	Asymmetric	An expected causal relation between Arg1 and $\neg$ Arg2 is cancelled or denied by Arg2.
7.	Contrast	Symmetric	One or more differences between Arg1 and Arg2 are highlighted with respect to what each predicates as a whole or to some entities they mention.
8.	Exception	Asymmetric	Arg2 indicates one or more circumstances in which the situation(s) described by Arg1 does not hold.
9.	Similarity	Symmetric	One or more similarities between Arg1 and Arg2 are highlighted with respect to what each predicates as a whole or to some entities they mention.
10.	Substitution	Asymmetric	Arg1 and Arg2 are alternatives, with Arg2 being the favoured or chosen alternative.
11.	Conjunction	Symmetric	Arg1 and Arg2 bear the same relation to some other situation evoked in discourse. Their conjunction indicates that they both hold with respect to that situation.
12.	Disjunction	Symmetric	Arg1 and Arg2 bear the same relation to some other situation evoked in the discourse, explicitly or implicitly. Their disjunction indicates that they are alternatives with respect to that situation, with the disjunction being non-exclusive so that both Arg1 and Arg2 may hold.
13.	Exemplification	Asymmetric	Arg1 describes a set of situations; Arg2 describes an element of that set.
14.	Elaboration	Asymmetric	Arg1 and Arg2 are the same situation, but Arg2 contains more detail.
15.	Restatement	Symmetric	Arg1 and Arg2 describe the same situation, but from different perspectives.
16.	Synchrony	Symmetric	Some degree of temporal overlap exists between Arg1 and Arg2. All forms of overlap are included.
17.	Asynchrony	Asymmetric	Arg1 temporally precedes Arg2.
18.	Expansion	Asymmetric	Arg2 is a situation involving some entity/entities in Arg1, expanding the narrative forward of which Arg1 is a part, or expanding on the setting relevant for interpreting Arg1. The Arg1 and Arg2 situations are distinct.
19.	Functional dependence	Asymmetric	Arg2 is a dialogue act with a responsive communicative function; Arg1 is the dialogue act(s) that Arg2 responds to.
20.	Feedback dependence	Asymmetric	Arg2 is a feedback act that provides or elicits information about the understanding or evaluation by one of the dialogue participants of Arg1.

Fonte: (ISO 24617-8, 2016, pp. 12-13)

Como se pode observar na tabela 4, todas as relações discursivas definidas pela ISO 24617-8 assentam na existência, ou não, de uma relação de dependência entre dois argumentos. Com base neste critério, a norma classifica as vinte relações discursivas normalizadas em dois grandes grupos: quatro relações simétricas e dezasseis relações

assimétricas.

Nas relações simétricas, os dois argumentos desempenham papéis equivalentes e podem ser permutados sem alteração do sentido global da relação. Por outras palavras, se uma relação se estabelece entre A e B, então a relação inversa entre B e A é semanticamente equivalente. Já nas relações assimétricas, os argumentos não são permutáveis: cada um desempenha um papel distinto e funcionalmente diferenciado, pelo que uma relação (A, B) não equivale a (B, A).

Importa sublinhar que o conjunto das vinte relações discursivas definidas na ISO 24617-8 não é fechado nem rígido. A norma prevê explicitamente a sua expansão por três vias principais: (i) através de refinamentos mais finos das relações nucleares; (ii) por meio de abstrações mais gerais dessas mesmas relações; e (iii) pela introdução de relações com o mesmo grau de granularidade que complementem o núcleo relacional definido pela norma (ISO 24617-8, 2016).

Para além do enquadramento teórico anteriormente referido, a ISO 24617-8 baseia-se igualmente na experiência acumulada na anotação de relações discursivas em múltiplas línguas e géneros textuais. Este trabalho empírico conduziu à definição de nomes mnemónicos para os papéis argumentais das relações discursivas assimétricas, apresentados na Tabela 5. Uma vez que estas relações envolvem argumentos com funções semânticas distintas, cada relação assimétrica especifica dois papéis argumentais diferenciados.

Tabela 5 - Nomes mnemónicos para os papéis argumentais das relações discursivas assimétricas

Relation	Arg1	Arg2
Cause	Result	Reason
Condition	Consequent	Antecedent
Negative Condition	Consequent	Negated Antecedent
Purpose	Enablement	Goal
Manner	Achievement	Means
Concession	Expectation-raiser	Expectation-denier
Exception	Regular	Exclusion
Substitution	Disfavoured alternative	Favoured alternative
Exemplification	Set	Instance
Elaboration	Broad	Specific
Asynchrony	Before	After
Expansion	Narrative	Expander
Functional Dependence	Antecedent act	Dependent act
Feedback Dependence	Feedback scope	Feedback act

Fonte: (ISO 24617-8, 2016, pp.13-14)

A título ilustrativo, apresentam-se de seguida quatro exemplos de relações causais em que, de forma consistente, o Arg2, com o papel de Razão, fornece uma explicação para o Arg1, que corresponde ao Resultado:

(42)

- a) Perhaps because **they won**, Mr. Bork's attackers come through more vividly than his defenders.
- b) Sears is negotiating to refinance its Sears Tower for close to \$850 million, sources said. (Implicit = because) **The retailer was unable to find a buyer for the building.**
- c) **Now, though, enormous costs for earthquake relief will pile on top of outstanding costs for hurricane relief.** "That obviously means that we won't have enough for all of the emergencies that are now facing us, and we will have to consider appropriate requests for follow-on funding," Mr. Fitzwater said.
- d) d) [+PA] The nations of Southern Africa know a lot about managing elephants; (Implicit = as) **their herds are thriving.**

A ISO 24617-8 procede ainda a um mapeamento sistemático das principais teorias de relações discursivas desenvolvidas até à data da sua publicação, evidenciando correspondências e convergências entre as relações propostas e as relações propostas pelos diferentes enquadramentos teóricos. Este mapeamento resulta de uma comparação detalhada das definições e dos critérios semânticos adotados em cada modelo.

Na Tabela 6, apresentamos apenas o mapeamento referente às relações causais, que constituem o objeto de estudo desta investigação.

Tabela 6 - Mapeamento de relação de causa (ISO 24617-8, 2016)

Teorias	Correspondentes à relação de Causa (ISO 24617-8)
RST	<i>Volitional cause, Non-volitional cause, Volitional result, Non-volitional result, Evidence, Justify</i>
RST Treebank	<i>Cause, Consequence, Result, Evidence, Explanation-argumentative, Reason</i>
HTDC	<i>Explanation, Cause, Evaluation</i>
GraphBank	<i>Cause-effect, Elaboration</i>
PDTB	<i>Reason, Result, Justification</i>
CCR/DiscAn	<i>Causal-Semantic-Basic-Positive, Causal-Semantic-NonBasic-Positive, Causal-Pragmatic-Basic-Positive, Causal-Pragmatic-NonBasic-Positive</i>
SDRT	<i>Explanation, Result</i>
DISCOR	<i>Explanation, Result</i>
ANNODIS	<i>Explanation, Result</i>

Fonte: (ISO 24617-8, 2016, pp. 28-30)

Este panorama comparativo evidencia a diversidade de abordagens à causalidade discursiva, ao mesmo tempo que sublinha o esforço de convergência e normalização promovido pela ISO 24617-8.

2.7. Conclusão

Neste capítulo procedeu-se a uma caracterização geral das principais propostas teóricas

no domínio das relações discursivas, acompanhada de uma revisão mais aprofundada da terminologia, das definições e das propriedades associadas às relações de causalidade.

A análise da literatura revelou que a formalização da causalidade no quadro de teorias de relações discursivas ou retóricas permite não só uma descrição mais sistemática e rigorosa das relações causais, como também uma representação semanticamente fundamentada do significado dos enunciados em que essas relações ocorrem.

Embora o conceito nuclear de causalidade seja transversal a todas as abordagens examinadas, a revisão evidenciou igualmente uma considerável heterogeneidade entre os diferentes quadros teóricos, quer no que diz respeito à classificação das relações causais, quer quanto aos critérios utilizados para a sua definição e representação. É neste contexto que a norma ISO 24617-8 se destaca como uma alternativa particularmente robusta. Com efeito, a sua abordagem abrangente às relações discursivas de nível local, aliada ao mapeamento sistemático entre diferentes tradições teóricas e à definição explícita da relação de Causa em termos dos papéis dos argumentos de Razão e Resultado, torna este enquadramento especialmente adequado para a investigação da causalidade discursiva.

Por estas razões, a análise dos dados apresentada no capítulo seis adotará como base a proposta da ISO 24617-8.

3. A inferência de causalidade no discurso: marcas linguísticas

As relações discursivas ou retóricas constituem um instrumento fundamental para descrever a estrutura de sequências de palavras, frases simples e complexas e parágrafos que compõem um texto, quer oral quer escrito (Mann & Thompson, 1987; Hobbs, 1985). Neste quadro, os elementos e as estruturas linguísticas assumem particular relevância, na medida em que contribuem para a inferência das relações de sentido que se estabelecem entre as unidades discursivas.

Este capítulo incide sobre os diversos elementos e estruturas linguísticas a partir dos quais se pode inferir a relação de causalidade num determinado discurso, em particular no discurso escrito. Como tem sido reiteradamente salientado, um discurso configura-se como um conjunto de ideias ou de orações organizadas e articuladas, no qual se estabelecem múltiplas relações discursivas. Estas relações podem ser identificadas através de processos de inferência que são, em certa medida, orientados por diferentes elementos gramaticais.

Segundo Van den Broek (1989), a compreensão e a interpretação de um discurso exigem, para além da competência linguística do falante, que as relações entre os eventos presentes no discurso sejam coesas e coerentes. Neste contexto, o processo de inferência das relações discursivas implica que o leitor mobilize conhecimento do mundo real e a capacidade de analisar aspetos específicos de determinados contextos, sobretudo nos casos implícitos, que tendem a ser mais difíceis de identificar no discurso. É, assim, neste domínio que se situa o cerne do presente capítulo, que se desenvolve a partir da análise das marcas linguísticas que expressam a ideia de causalidade, isto é, que marcam as relações discursivas de causa.

Neste sentido, importa lembrar que as construções causativas, na língua inglesa, se apresentam sob duas formas, segundo Girju (2003), relações explícitas e relações implícitas. De acordo com o autor,

explicit causation patterns can contain relevant keywords such as *cause*, *effect*, *consequence*, but also ambiguous ones such as *generate*, *induce*, etc. The implicit causative constructions are more complex, involving inference based on semantic analysis and background knowledge. (Girju, 2003, p. 3).

As construções causativas explícitas envolvem, em geral, expressões linguísticas

que assinalam diretamente a existência da relação causal. Nesse sentido, Khoo et al. (2002) argumentam que:

many of the cause-effect relations in text are implied and have to be inferred by the reader, there is also a wide variety of linguistic expressions for explicitly indicating cause and effect. In addition, it has been found that certain words have “causal valence” – they bias the reader to attribute cause in certain ways (Khoo et al., 2002, p. 1).

Efetivamente, vários autores têm sublinhado o papel dos elementos ou marcas linguísticas na introdução e articulação das ideias que estruturam um discurso. Entre esses autores destacam-se Arrais (1985), Asher & Lascarides (2003), os estudos desenvolvidos por Prasad et al. (2008), Silvano (2010) e ISO 24617-8 (2016), todos eles referidos nos capítulos anteriores.

Neste enquadramento, é consensual a afirmação de que um aspecto crucial da compreensão e da produção do discurso depende do reconhecimento de diferentes relações discursivas presentes em textos. Nesse contexto, os conetores discursivos podem sinalizar essas relações, o que entendidos como predicados discursivos (Prasad et al., 2010). De acordo com Couper-Kuhlen e Kortmann (2000, p. 1), “these relations can be realized or marked by different linguistic means, e.g. by adverbials, particles, coordinating and subordinating conjunctions, word order”.

De facto, a inferência das diversas relações discursivas pode apoiar-se em vários elementos linguísticos, nomeadamente nomes, verbos, advérbios, preposições, os conetores e construções como orações coordenadas e subordinadas. Estes elementos e estruturas serão desenvolvidos de forma mais detalhada nos pontos que se seguem, com especial incidência no léxico, preposições, conetores discursivos e nas orações, no âmbito específico das relações discursivas de causalidade.

3.1. Léxico

No processo de produção e de interpretação textual constrói-se uma harmonia de sentido que resulta das conexões estabelecidas entre palavras e expressões, configurando uma “rede de relações semânticas – estabelecidas no interior de um texto – resultante da exploração de lexemas que partilham, total ou parcialmente, traços semânticos” (Lopes, 2013, p. 35). Estas conexões não operam apenas ao nível das frases

isoladas, mas contribuem decisivamente para a articulação do discurso e para a inferência das relações que nele se estabelecem.

Neste enquadramento, o léxico assume um papel central. Um conjunto de palavras organizado em torno de um campo conceptual comum estabelece relações de significado que permitem estruturar a interpretação textual. No léxico de cada língua, entendido como o vasto conjunto de palavras que a constitui, encontram-se inúmeros campos lexicais, que correspondem a subconjuntos de palavras associadas a uma mesma área de conhecimento. É precisamente a ativação destes campos que contribui para a construção de coerência discursiva.

De acordo com Asher & Lascarides (2003), os itens lexicais pertencentes a classes gramaticais como nomes comuns, verbos, adjetivos, entre outros, veiculam informação relevante para a dedução de diferentes relações discursivas. Os mesmos autores sublinham que “the lexicon [is] to constitute the body of linguistic information concerning the syntactic and semantic relationships among words” (Asher & Lascarides, 2003, p. 250). Assim, o léxico não se limita a codificar significados isolados, mas integra um conjunto de informações que permite estabelecer relações entre unidades linguísticas.

O contributo do léxico para a interpretação textual manifesta-se em diferentes níveis. Por um lado, relaciona-se com as restrições convencionais de significado e de uso das palavras, sendo fundamental para a interpretação correta de frases individuais. Por outro lado, desempenha um papel crucial no cálculo ou na computação das relações retóricas entre frases. Com efeito, as palavras adquirem sentidos específicos em função do contexto em que ocorrem, o que remete para o domínio do campo semântico lexical. Paralelamente, é fundamental distinguir o significado primário e objetivo da palavra, o seu sentido denotativo, o significado secundário e subjetivo, associado ao processo de inferência e ao valor conotativo.

Uma compreensão clara destas dimensões do léxico permite modelar a interação entre léxico e discurso, com base na sintaxe e na semântica composicional simples. Como é sabido, a semântica ocupa-se do estudo do significado linguístico expresso pelas unidades simples e pelas suas combinações, distinguindo-se entre semântica lexical,

centrada nas unidades simples, e semântica composicional, que incide sobre as expressões complexas. Neste sentido, o léxico consiste num conjunto de informações linguísticas que representam as relações sintáticas e semânticas entre palavras, sendo essencial para o cálculo das contribuições lexicais para a forma lógica.

Para além disso, o léxico articula-se com o conhecimento do mundo real e envolve inferências frequentemente consideradas autoevidentes. Nomes, verbos e adjetivos, enquanto diferentes tipos lexicais, denotam informações relativas a tipos semânticos de objetos e permitem inferir subtipos (A, B). Em particular, o léxico causal transmite informações sobre relações causais, podendo constituir uma pista determinante para a inferência de relações de causa entre eventos ou estados (Asher & Lascarides, 2003).

Como demonstram Asher & Lascarides (2003) e Silvano (2010), o conhecimento do léxico e o conteúdo proposicional de frases justapostas podem conduzir o leitor à inferência das relações discursivas, mesmo na ausência de marcadores explícitos:

(43) Max caiu. O João ajudou-o a levantar-se.

(44) Max caiu. John empurrou-o.

Nestes exemplos, não se verifica a presença de qualquer conetor explícito que defina a relação discursiva entre as frases. No entanto, com base no contexto e no conhecimento lexical, é possível inferir diferentes relações discursivas. No primeiro caso, estabelece-se uma relação de Narração entre “caiu” e “levantar-se”, que expressa uma sequência cronológica, indicando que a primeira situação ocorre antes da segunda. No segundo exemplo, evidencia-se uma relação de Explicação, sendo a causa inferida a partir do léxico e do conteúdo proposicional.

Uma análise semelhante pode ser aplicada ao exemplo seguinte (Silvano, 2010, p. 217):

(45) John comprou uma mala. Ele vai viajar.

Também aqui as situações expressas por “comprou” e “viajar” se relacionam sem a presença de um conector. Ainda assim, é possível inferir a existência de uma relação discursiva de explicação. Tal como refere Silvano (2010):

the lexicon, namely the fact that «suitcase» and “travel” belong to the same lexical domain, and our world knowledge that, whenever someone travels, a suitcase is normally needed leads to the interpretation that the second sentence represents the cause for the situation represented by the first sentence (Silvano, 2010, p. 218).

Deste modo, o conhecimento lexical revela-se essencial para a identificação do conteúdo proposicional e para a inferência das relações discursivas.

Neste subcapítulo, apresentaremos algumas considerações sobre elementos lexicais que podem contribuir para a inferências das relações causais: os nomes causativos, os verbos causativos e os advérbios.

3.1.1. Verbos causativos

Um dos elementos lexicais mais determinantes e mais estudados para as relações causais são os verbos causativos, que se caracterizam por presença de um elemento causal na sua estrutura semântica, como os verbos *matar*, *magoar*, *quebrar*. Estes verbos são, em geral, transitivos, embora, dependendo do contexto, possam admitir usos intransitivos. Por exemplo, os verbos *matar* e *quebrar* exigem tipicamente um complemento, como se observa nas questões *Quem mata, mata quem?* e *Quem quebra, quebra o quê?*.

Os verbos causativos podem ser subdivididos em três categorias, conforme proposto por Girju (2003):

1. Simple causatives (cause, lead to, bring about, generate, make, force, allow, etc.) Here the linking verb refers only to the causal link, being synonymous with the verb cause. E.g., “Earthquakes generate tidal waves.”;
2. Resultative causatives (kill, melt, dry, etc.) These verbs refer to the causal link plus a part of the resulting situation;
3. Instrumental causatives (poison (killing by poisoning), hang, punch, clean, etc.) These causatives express a part of the causing event as well as the result.

(Girju, 2003, p. 2)

De igual modo, Thompson (1987) (apud por Khoo et al. (2002)), classificou os verbos causativos em três grupos distintos:

1. Transitive causative verbs that also have an intransitive usage. For example, “x breaks y” is paraphrased as “x causes y to break.”;
2. Causative verbs that do not have an intransitive usage. The transitive kill is paraphrased using the intransitive die, a different word: “x kills y” is paraphrased as “x causes y to die.”;
3. Causative verbs that do not have any intransitive verb that can be used in the paraphrase. The past participle form of the causal verb is used instead. “X butters y” is paraphrased as “x causes y to be buttered.”

(Thompson, 1987, apud por Khoo, Chan, & Niu, 2002, p. 6).

Como refere Thompson (1987), todos os verbos transitivos, em particular os verbos de ação, podem ser considerados causativos, partindo do princípio de que toda a ação tende a ter uma causa que explica ou justifica a sua ocorrência. Tal como ilustrado na paráfrase “x causes y to break”, esta relação pode ser observada nos exemplos seguintes:

(46) A queda da madeira quebrou a parede do cercado.

(47) A parede do cercado quebrou.

Importa, contudo, salientar que estas classificações não são consensuais, existindo controvérsias quanto à distinção entre verbos causativos e outros verbos transitivos de ação. Alguns verbos, como *bater*, *chutar* ou *morder*, não devem ser equiparados a verbos causativos, apesar de expressarem ações. Estes verbos são tipicamente associados a agentes humanos e não aceitam eventos ou estados de coisas como sujeitos, ao contrário dos verbos causativos. Apesar destas questões, Khoo et al. (2002, p. 6) apresentam exemplos que permitem clarificar a distinção entre verbos causativos e outros verbos de ação:

(48) Oswald killed Kennedy by shooting at him.

(49) Oswald's accurate shooting killed Kennedy.

(50) John broke the vase by shooting at it.

(51) John's accurate shooting broke the vase.

(52) Tom hit the can by shooting at it.

(53) *Tom's accurate shooting hit the can.

A partir destes exemplos, os autores explicam que os exemplos (48) e (50) demonstram que o sujeito dos verbos *kill* e *break* pode ser um evento, ao passo que os exemplos (51) e (52) mostram que o sujeito do verbo *hit* tem de ser uma pessoa ou um objeto. Segundo Thompson (1987, p. 6), isto ocorre porque, nos verbos de ação como *hit*, o sujeito não é separável do resultado da ação.

Com base nestas observações, e apoiando-se nas propostas de Thompson (1987) e Szeto (1988) e Khoo et al. (2002) propõem três critérios para distinguir verbos causativos de outros verbos transitivos de ação:

1. Causative verbs accept events and states of affairs as subjects, whereas other action verbs accept only agents as subjects (Thompson, 1987).
2. Causative verbs are transitive verbs which also have an intransitive usage where the subject of the verb has the patient role (Szeto, 1988).
3. Causative verbs specify the result of the action, whereas other action verbs specify the action but not the result of the action (Szeto, 1988).

(Khoo et al., 2002, p. 7)

Uma outra contribuição relevante para o estudo dos verbos causativos é a classificação proposta por Levin (1993), que defende que existe uma correlação sistemática entre o significado lexical dos verbos e o seu comportamento sintático. Este princípio é particularmente relevante para a análise dos verbos causativos, uma vez que a causalidade se manifesta frequentemente através de alternâncias na realização dos argumentos.

Um fenómeno central é a alternância causativa/incoativa, que envolve verbos que ocorrem tanto em construções transitivas como intransitivas. Na variante transitiva, o evento é apresentado como causado por um argumento externo, enquanto na variante intransitiva se expressa apenas a mudança de estado do participante afetado. Levin (1993) mostra que esta alternância é característica sobretudo dos verbos de mudança de estado. No entanto, nem todos os verbos semanticamente causativos admitem esta alternância, uma vez que a sua disponibilidade depende de restrições semânticas específicas.

De acordo com Soares da Silva (1999, p. 550), “entende-se geralmente por causativo o verbo que incorpora a noção de causa-causalidade na sua significação, sendo

por isso parafraseável por causar uma acção ou causar a aquisição de uma qualidade ou a entrada num estado”. O autor considera que as construções verbais causais podem ser formadas através de três modos diferentes:

Causativos lexicais: expressam a causalidade num único verbo simples (exs.: matar, abrir, mostrar, ensinar, ferver);

Causativos morfológicos: expressam a causalidade num verbo derivado (nem sempre derivado "lexical"), mais precisamente, o evento-causador é expresso (geralmente) por um afixo (ou mais) e o evento-causado/ efeito pela forma simples (amarelecer, adormecer, arrefecer, humanizar, purificar, endireitar, secar, enrugar, dificultar);

Causativos analíticos: expressam a causalidade através de uma construção sintáctica formada por dois predicados, em que um exprime o evento-causador (fazer, mandar, deixar, pôr) e o outro o evento-causado (estudar / a estudar); ou o que muitos autores não referem — uma construção formada por um predicado e um elemento predicativo (fazer/pôr/tornar/deixar feliz).

(Soares da Silva, 1999, pp. 554-555)

Independentemente do seu processo de derivação ou composição, os verbos causativos são um elemento central na expressão da causalidade no discurso.

3.1.2. Nomes causativos

Os nomes podem igualmente marcar relações causais ou incorporar, à semelhança dos verbos, elementos de natureza causal. Incluem-se aqui nomes lexicais como *causa*, *motivo* ou *razão*, mas também nomes de verbais como *destruição*, *explosão* ou *inundação*, que codificam eventos de causação ou os seus resultados. Estes nomes assumem particular relevância na descrição semântica, uma vez que permitem representar relações causais fora do domínio estritamente verbal. Esta secção incidirá nas nominalizações, que têm sido objeto de investigação sistemática em diversos tipos de abordagens, lexicais, semânticas e sintáticas.

No caso do português europeu, Brito & Oliveira (1997) argumentam que as nominalizações deverbais podem apresentar três leituras distintas: uma leitura de evento, uma leitura de resultado e uma leitura referencial ou individual. As autoras mostram ainda que, quando derivadas de verbos de realização, estas nominalizações tendem a preservar a natureza aspetual do verbo correspondente, a menos que ocorra uma mudança para uma leitura individual. Assim, em exemplos como *A Maria está a fazer a encomenda de livros*, a nominalização *encomenda* recebe uma leitura eventiva;

em *A encomenda de livros enriqueceu a nossa biblioteca*, a mesma forma é interpretada como resultado; e em *Vou mandar esta encomenda*, assume uma leitura referencial, designando um objeto individual.

Apesar das divergências teóricas, existe consenso quanto ao facto de certas nominalizações preservarem propriedades associadas ao verbo de base. Grimshaw (1990) propõe, numa perspetiva lexicalista, uma distinção fundamental entre nominais de evento complexo, que incluem um parâmetro de evento na sua estrutura argumental, e nominais de resultado, que não incluem esse parâmetro.

De um ponto de vista semântico, os nomes causativos são, em geral, nomes deverbais que denotam eventos de causação ou os resultados produzidos por esses eventos, como *explosão* (de *explodir*) ou *destruição* (de *destruir*). Estas nominalizações, assim como outras sem valor causal, ocupam um lugar central no léxico por exibirem frequentemente uma polissemia regular entre uma leitura eventiva, correspondente ao processo ou ato, e uma leitura resultativa, correspondente ao efeito ou estado resultante (Pustejovsky, 1995). Um dos principais desafios teóricos consiste precisamente em explicar como um mesmo nome pode alternar sistematicamente entre estas duas leituras e quais os mecanismos linguísticos que desencadeiam cada interpretação.

Diversos estudos mostram que as nominalizações, nomeadamente as causativas, tendem a admitir, pelo menos, duas leituras principais. Na leitura de evento, o nome denota o ato ou processo, como em *a destruição da cidade*, interpretada como “o ato de destruir a cidade”. Na leitura de resultado, o mesmo nome pode denotar o produto ou estado resultante desse ato, como em *a destruição da cidade* entendida como o conjunto de escombros ou o estado de devastação. Grimshaw (1990) identifica um critério sintático relevante associado a esta alternância: a presença de um complemento interno favorece fortemente a leitura eventiva, enquanto a sua ausência tende a favorecer a leitura resultativa. Este contraste é ilustrado, por exemplo, em inglês pela oposição entre *the examination of the patients*, que só admite uma leitura eventiva, e *the examination was on the table*, que apenas permite uma leitura de resultado.

Um traço central destas nominalizações com leitura eventiva é a preservação da estrutura argumental do verbo de base, de forma semelhante ao que ocorre com os

verbos em contexto frásico (Grimshaw, 1990; Alexiadou, 2001). Em termos temáticos, o argumento interno do verbo, geralmente o paciente ou tema afetado, permanece disponível e surge frequentemente realizado como complemento nominal, como em *destruição da cidade*, na qual *cidade* corresponde ao argumento afetado. O argumento externo, correspondente ao agente ou causa, pode igualmente ser expresso, quando pertinente, sob a forma de um adjunto, tipicamente introduzido pela preposição *por*, como em *a destruição da cidade pelo exército*.

Importa sublinhar, contudo, que os nomes, ao contrário dos verbos, não impõem a realização obrigatória dos seus argumentos. Os argumentos nominais são opcionais, sendo perfeitamente gramatical a ocorrência de formas como *a destruição*, com agente e tema implícitos (Borer, 2003). Em contraste, a frase verbal correspondente exige, pelo menos, a realização do objeto interno. Ainda assim, quando mais do que um argumento é explicitamente realizado, como em *a destruição da cidade pelo exército*, a leitura eventiva torna-se inequívoca, uma vez que os nomes de resultado não licenciam, em geral, a expressão simultânea de múltiplos argumentos verbais.

A agentividade desempenha igualmente um papel determinante na interpretação das nominalizações, inclusive das causativas. Nomes deverbais derivados de verbos transitivos agentivos, como *destruir* ou *assassinar*, implicam tipicamente a presença de um agente causal, expresso ou implícito. Assim, *o assassinato de César* pressupõe necessariamente a existência de um agente humano. Em contraste, nominalizações derivadas de verbos não-agentivos ou incoativos, como *morte*, não implicam obrigatoriamente um agente, podendo denotar eventos espontâneos ou causados por fatores naturais. De forma semelhante, nominalizações de verbos inacusativos não admitem a introdução de um agente externo. Já nomes como *explosão* permitem tanto uma leitura não-agentiva, como em *a explosão do foguete*, equivalente a “o foguete explodiu”, como uma leitura agentiva, quando o verbo de base admite uma estrutura causativa, como em *a explosão da ponte pelo exército*. Em suma, a estrutura temática herdada do verbo de base condiciona os participantes que podem ser expressos pelo nome e influencia decisivamente a leitura semântica obtida (evento vs. resultado).

Apesar da ambiguidade inerente a muitas nominalizações, o contexto linguístico tende, na maioria dos casos, a desambiguar a interpretação, frequentemente através de processos de coerção semântica. Certos modificadores ou construções favorecem uma leitura específica: adjetivos agentivos ou aspetuais tendem a selecionar a leitura eventiva, como em *a destruição deliberada da cidade* ou *a destruição gradual da cidade*, que pressupõem um processo dinâmico. Por outro lado, modificadores que descrevem estados finais favorecem a leitura de resultado, como em *a destruição completa da cidade*. De modo semelhante, alguns predicados verbais selecionam complementos nominais de tipo eventivo, como *ocorrer*, *acontecer* ou *ter lugar*, enquanto predicados estativos ou locativos podem induzir uma leitura resultativa. Assim, contrastes como *A construção do edifício durou dois anos* versus *A construção do edifício fica na colina* mostram que o mesmo nome pode receber interpretações distintas em função das propriedades sintático-semânticas do contexto (Brito, 2003).

3.1.3. Advérbios com valor causal

Quanto aos valores causais introduzidos por advérbios, destaca-se o estudo de Cresswell (1981), no qual se defende que o significado de certos advérbios em inglês integra um elemento causal. Um advérbio causal simples, como *fatally*, cujo significado implica que ocorre uma morte causada por um evento, pode expressar tanto “the result of an action” como “a way that it is the subject who dies” (Cresswell, 1981, p. 23). Esta distinção é ilustrada pelos exemplos seguintes:

(54) Arabella fatally wounded Bramwell.

(55) Catherine fatally slipped.

(Cresswell, 1981, p. 23)

No exemplo (54), entende-se que a ação de Arabella causa a morte de Bramwell, enquanto, no exemplo (55), o escorregar de Catherine é interpretado como a causa da sua própria morte. Em ambos os casos, identifica-se uma relação causal marcada pelo advérbio *fatally*.

Para além deste tipo de advérbios, Cresswell (1981) distingue ainda seis situações

em que os advérbios expressam valores causais. Em primeiro lugar, os advérbios de percepção, nos quais uma ação provoca ou gera um efeito perceptível, ou a percepção de alguém, como *audibly* e *visibly*, como se pode observar no exemplo seguinte. Nestes casos, a relação causal restringe-se a uma causalidade facultativa ou direta, na medida em que a percepção só ocorre após a realização do evento.

(56) {Fred, (sings, audibly)}

(Cresswell, 1981, p. 26)

Em segundo lugar, encontram-se os advérbios menos especificamente perceptivos, ou marginalmente perceptivos, como *manifestly*, *patently*, *publicly* e *conspicuously*, que remetem para resultados semelhantes aos perceptivos, embora de forma menos direta, como se pode observar no exemplo seguinte.

(57) O presidente foi publicamente julgado, o que levou à sua demissão na segunda-feira.

Em terceiro lugar, surgem os advérbios causais com resultados dependentes do contexto. Deste grupo fazem parte advérbios que introduzem resultados esperados ou desejados decorrentes da ação causal, resultados esses que se ajustam às circunstâncias em que ocorrem. Inserem-se neste grupo advérbios como *successfully*, *plausibly*, *conveniently*, *amusingly* e *pleasantly*, como se mostra no seguinte exemplo.

(58) A menina reagiu agradavelmente aos professores, facilitando a interação.

Em quarto lugar, identificam-se os advérbios causais com resultados possíveis, em que uma situação, seja um evento ou um estado, apresenta a possibilidade de produzir um determinado efeito, sem que tal efeito seja necessariamente garantido no mundo real. Exemplos deste tipo são *irrevocably*, *tenuously*, *precariously* e *rudely*. Como se pode observar no exemplo seguinte.

(59) O acordo foi estabelecido tenuously, dependendo de fatores externos instáveis.

Em quinto lugar, Cresswell refere os advérbios focados no efeito, e não na causa, que salientam o resultado ou o efeito originado por um dado anterior, como *gratefully*, *consequently* e *painfully*. Este valor é também reconhecido por Castilho (2014), que, tal como outros autores, integra o advérbio *consequentemente* no grupo dos conetores. Como se observa no exemplo seguinte, este advérbio estabelece uma relação de causa e consequência:

(60) Traduzir é servir. Consequentemente, trabalho de inferiores. (Castilho 2014, p. 581)

Por fim, o autor identifica os advérbios causais com valor de meio, nos quais os advérbios introduzem o próprio processo causal, destacando a maneira pela qual o resultado é alcançado, como em *mechanically* ou *magically*. Como se pode observar no exemplo seguinte.

(61) O motor desligou-se mecanicamente após o sobreaquecimento.

Considerando que os advérbios podem modificar o significado dos adjetivos, Cresswell (1981, p. 30) afirma que os “adverbs (...) permit a closely related semantics for the corresponding adjectives”. Deste modo, os valores dos advérbios causais permitem inferir que os adjetivos correspondentes também podem implicar uma relação causal, em que uma situação provoca outra. Para ilustrar esta relação, o autor compara as formas *fatal* e *fatally*. Ambas, quer na forma adjetival quer na forma adverbial, podem denotar causalidade; contudo, existe uma diferença semântica entre elas. A forma adjetival tende a focar-se no evento externo que causa outro evento, enquanto a forma adverbial destaca a maneira como o evento ocorreu e a própria ação. Esta diferença pode ser observada nos exemplos seguintes:

(62) Guinevere's fatal walk.

(63) Guinevere walked fatally.

(Cresswell, 1981, p. 31)

3.2. Preposições

As preposições podem igualmente funcionar como pistas para a inferência de relações discursivas, em particular das relações causais.

Um contributo relevante para o estudo das preposições com valor causal é o de Dirven (1995), que analisa a expressão da causa a partir de uma perspetiva cognitiva. Segundo o autor, a causalidade não constitui uma relação homogénea ou unidimensional, sendo antes concetualizada de formas diversas, frequentemente ancoradas em esquemas espaciais e experiencialmente motivadas. Dirven mostra que várias preposições do inglês, cuja função primária é espacial, podem adquirir leituras causais em função do modo como o falante concetualiza a relação entre dois eventos ou estados de coisas. Em particular, o autor analisa preposições que exprimem relações de proximidade, de origem e de envolvimento ou contenção, evidenciando que estes esquemas espaciais servem de base para a interpretação causal.

No domínio da proximidade conceptual, preposições como *at*, *with*, e *by*, permitem construir a causalidade a partir da contiguidade ou da interação entre duas situações, como ilustram os exemplos seguintes:

(64) He was irritated at her request.

(65) He felt irritated with her and wished she would go.

(66) He was irritated still further by her tone.

Dirven (1995, p. 105)

Noutros casos, a causalidade é concetualizada como origem, sendo expressa por preposições que remetem para um ponto, uma linha ou uma área de proveniência, como *from*, *of* ou *out of*, em expressões como *die from alcohol*, *die of anemia* ou *act out of love*. Nestes contextos, a causa é entendida como a fonte a partir da qual emerge o evento consequente. Por fim, Dirven analisa preposições que concetualizam a causa

como um tipo de envolvimento ou de “volume” concetual, em que o evento causador é entendido como um espaço abstrato que circunscreve o estado ou a ação descritos. É o caso de preposições como *in*, *about* e *over*, como se observa nos exemplos seguintes:

(67) But I really do not find much pleasure in going over manufactories.

(68) I am delighted about my promotion.

(69) Denmark has rejected a protest over an agreement with the Baltic republics.

No âmbito do português, Cunha & Cintra (1984) referem que as preposições podem estabelecer relações entre palavras, exprimindo valores semânticos temporais, espaciais ou nocionais. Em certos casos, as leituras nocionais podem veicular valores causais, como sucede com a preposição *de*, conforme ilustrado no exemplo seguinte:

(70) Chorava de dor. (Cunha & Cintra, 1984, p. 553)

Neste caso, a preposição *de* estabelece uma relação nocional entre o antecedente *chorava* e o conseqüente *dor*, sendo este último interpretado como explicação do primeiro, o que permite identificar uma relação causal entre os dois termos.

Considerando a capacidade das preposições para estabelecer relações entre elementos de uma mesma frase, Brito (2003) descreve os valores sintático-semânticos de algumas preposições particularmente relevantes para a expressão da causalidade. Entre estas, destacam-se:

- *por*, que pode exprimir os papéis temáticos de agente, instrumento, causa ou razão (pp. 396-397), como se observa nos exemplos:

(71) A cidade foi destruída pelo exército/ pela bomba.

(72) A destruição da cidade pelo exército/ pela bomba.

(73) Por ela sou capaz de tudo.

- *de*, que pode introduzir certos complementos verbais com valor de razão ou causa (Brito, 2003, p. 397), como em:

(74) Eles choraram de dor.

- *com*, que pode exprimir valores de causa ou razão (p. 398), como ilustrado em:

(75) Com o temporal, as aulas não puderam começar.

Um contributo mais recente é o de Raposo & Xavier (2020), que analisam os valores semânticos de várias preposições associadas à causalidade, nomeadamente *de*, *por*, *com*, *visto* e *dado*. Segundo estes autores, a preposição *de*, para além de remeter para uma leitura espacial dinâmica, pode igualmente marcar valores causais, como se observa em:

(76) Todos festejaram de alegria. (Raposo & Xavier 2020, p. 1550)

A preposição *por* introduz frequentemente o agente da passiva, que corresponde ao sujeito da oração ativa e pode desempenhar o papel temático de causa, frequentemente associado a uma força não animada com propriedades causais, como em:

(77) Aquela casa foi destruída por um incêndio. (Raposo & Xavier 2020, p. 1554),

Além disso, *por* pode igualmente introduzir constituintes com valor causal em orações não passivas, como se observa em:

(78) Por esse motivo, não volto a falar contigo. (Raposo & Xavier 2020, p. 1554)

A preposição *com* introduz valores mais específicos, como modo, causa,

comitativa e instrumental. De acordo com Raposo & Xavier (2020), esta preposição introduz “expressões com o papel temático de causa, que representam forças (frequentemente, mas não necessariamente, não humanas) que têm uma atividade causal não volitiva e não consciente” (Raposo & Xavier, 2020, p. 1558), como em:

(79) O telhado voou com o furacão.

Para além disso, *com* pode selecionar um complemento de núcleo adjetival, nominal ou preposicional, introduzindo um adjunto adverbial com o sentido de dada a situação ou por causa da situação, como em:

(80) Com as crianças doentes, a Maria ficou em casa. (Raposo & Xavier 2020, p. 1558)

Por fim, importa referir que locuções prepositivas com valor causal, como *por causa de* e *graças a*, não são descritas de forma sistemática em Cunha & Cintra (1984), ou Raposo & Xavier (2020). Contudo, estas locuções podem claramente ter um valor causal, como atestado em dicionários, como o Dicionário da Academia de Ciências de Lisboa.

3.3. Conectores

Do ponto de vista de Lopes & Carrilho (2020), o processo de estruturação e organização textual recorre a determinados elementos linguísticos que asseguram a articulação e a progressão ao longo da composição do texto. Estes elementos são fundamentais para a criação de coesão e coerência discursivas. Com efeito, no processo de interpretação, constrói-se uma representação mental que integra a informação proveniente dos diferentes enunciados numa totalidade significativa. Assim, a compreensão plena de um texto não se limita à interpretação de frases ou enunciados isolados, mas envolve igualmente as conexões discursivas, isto é, a forma como os diferentes enunciados se relacionam entre si no contexto do discurso. Neste enquadramento, os conectores discursivos desempenham a função de estabelecer pontes entre segmentos do discurso,

ancorando a informação nova em conhecimentos já dados no contexto linguístico e partilhados pelos interlocutores.

Quanto à terminologia utilizada para designar estes elementos, a diversidade de abordagens linguísticas, assentes em diferentes quadros teóricos, deu origem à coexistência de várias designações concorrentes, tais como: frases de sinalização (Knott & Dale, 1994), conectivos discursivos (Blakemore, 1987), partículas discursivas (Schorup, 1985), conectivos fáticos (Bazanella, 1990), conectivos pragmáticos (Van Dijk, 1979; Stubbs, 1983), marcadores pragmáticos (Fraser, 1990; Schiffrin, 1987) e *semantic conjuncts* (Quirk et al., 1985). Apesar da diversidade terminológica, todas estas propostas convergem no reconhecimento do papel fundamental destes elementos linguísticos na garantia de uma articulação adequada entre diferentes unidades de sentido de um texto. Nesse sentido, Portolés (1993) define o conector da seguinte forma:

conector es un marcador discursivo que vincula semántica y pragmáticamente un miembro del discurso con otro miembro anterior. El significado del conector proporciona una serie de instrucciones que guían las inferencias que se han de obtener del conjunto de los dos miembros relacionados” (Apud, Martín Zorraquino & Portolés Lázaro 1999, p. 4093).

Já Fraser (1999, p. 950) adota a perspetiva de que os marcadores discursivos constituem “a pragmatic class, lexical expressions drawn from the syntactic classes of conjunctions, adverbials, and prepositional phrases”. Lopes & Carrilho (2020) definem os marcadores discursivos também em termos pragmáticos considerando a função conetiva na construção do texto e procedem igualmente a uma caracterização destes elementos nos planos sintático, morfológico e semântico. Do ponto de vista sintático, estes elementos articulam, em regra, enunciados adjacentes, podendo também estabelecer uma ligação entre o enunciado que introduzem e um segmento discursivo mais amplo, constituído por uma sequência de enunciados anteriores. Em termos morfológicos, são unidades invariáveis, uma vez que não admitem qualquer tipo de flexão, e, quando são formadas por mais do que um item lexical, funcionam como sintagmas totalmente cristalizados. Do ponto de vista semântico, não contribuem para o conteúdo proposicional do enunciado que tipicamente introduzem, sendo o seu significado de natureza procedimental, na medida em que codificam instruções relativas

à articulação entre enunciados, desempenhando, assim, uma função orientadora no processo interpretativo.

Em termos de distribuição textual, a sua ocorrência é bastante variável, ao contrário do que sucede com as conjunções de coordenação e de subordinação adverbial. Do ponto de vista sintático, Lopes & Carrilho (2020) salientam que estes elementos “não têm uma posição fixa, apresentando alguma mobilidade. A posição prototípica é a posição inicial, na periferia esquerda do segundo enunciado, mas podem ocupar outras posições” (p. 2682). Do ponto de vista morfológico e prosódico, as mesmas autoras referem que os marcadores discursivos são assinalados “por pausas, à direita e à esquerda, funcionando como constituintes prosódicos independentes, com um contorno entoacional parentético, assinalado na escrita por pontuação (MD entre vírgulas, MD demarcado à direita por ponto e vírgula ou ponto, e à esquerda por vírgula)” (p. 2683).

As autoras referem ainda que os marcadores discursivos apresentam polifuncionalidade sincrónica, podendo desempenhar múltiplas funções comunicativas ou semânticas, sobretudo no caso de formas homónimas. Além disso, não possuem uma função fixa e universal, uma vez que a sua interpretação pode variar consoante o contexto discursivo, sendo, por isso, fortemente dependentes desse contexto.

Etimologicamente, como refere Castilho (2014, p. 581), no processo de gramaticalização, alguns conetores derivaram de advérbios. Lopes & Carrilho (2020, p. 2684) sublinham igualmente que estas expressões, após um processo de reanálise, passam a desempenhar funções ao nível da organização textual.

advérbios (agora, assim, consequentemente, enfim, então, já, logo, ...) e de sintagmas preposicionais (daí, de outro modo, de qualquer modo, em particular, em primeiro lugar, ...) que, em estados de língua anteriores, desempenharam apenas funções sintáticas e semânticas no quadro da frase e que, entretanto, sofreram um processo de reanálise, passando a desempenhar também funções ao nível da organização textual (Lopes & Carrilho, 2020, p. 2684).

Este processo pode ser ilustrado pela oposição entre usos temporais e discursivos de advérbios como *então*, como se observa nos exemplos apresentados.

(81) Em 1974, houve uma revolução em Portugal. As pessoas tinham então uma enorme esperança de transformação do país

(82) Não há condições para levar a bom termo uma reestruturação profunda da oferta curricular. Então, resta-nos propor algumas alterações pontuais.

Em certos contextos, o advérbio mantém o seu valor temporal, noutros, funciona como marcador discursivo, estabelecendo uma relação causal ou conclusiva entre enunciados.

De acordo com Lopes & Carrilho (2020), as articulações que os marcadores discursivos estabelecem podem ocorrer “a nível local, entre enunciados adjacentes, e (...) mais global da estrutura temático-informacional e material do texto”, sendo que alguns sinalizam conexões de conteúdo entre enunciados, outros operam no domínio dos atos de fala e outros ainda no domínio da organização temático-informacional do texto (Lopes & Carrilho, 2020, p. 2684).

No estudo dos conetores na língua espanhola, Martín Zorraquino & Portolés Lázaro (1999) propõem uma classificação em três categorias: aditivos, consecutivos e contra-argumentativos.

O primeiro grupo estabelece a ligação entre um membro discursivo e outro anterior com a mesma orientação argumentativa (Martín Zorraquino & Portolés Lázaro, 1999). Este grupo subdivide-se em dois subtipos: os conetores que seguem uma escala argumentativa e reforçam a informação adicionada, hierarquizando a relação entre dois argumentos, como incluso, *inclusive* y *es más* (Martín Zorraquino & Portolés Lázaro, 1999) e os que não seguem uma escala argumentativa, limitando-se a introduzir informação adicional, como *além disso*.

A segunda categoria apresenta o membro do discurso como consequência de um membro anterior. Esta classe subdivide-se em quatro subtipos: os conectores *pues* e *así pues*, que se limitam a apresentar o segmento como consequente de um membro anterior os conectores *por tanto*, *por consiguiente*, *por ende* e *de ahí*, que fundamentam a passagem do antecedente ao consequente num raciocínio; os conectores *en consecuencia* e *de resultas*, nos quais o consequente corresponde a um estado de coisas

resultante de outro estado de coisas; e, por fim, as unidades menos gramaticalizadas como conectores consecutivos, tais como *así* e *entonces* (Martín Zorraquino & Portolés Lázaro, 1999).

A terceira categoria, contra-argumentativos, estabelece uma relação entre dois membros do discurso, em que o segundo se apresenta como supressor ou atenuador de uma conclusão potencialmente inferível a partir do primeiro. Fazem parte desta categoria marcadores discursivos como *pero*, *sin embargo*, entre outros.

De igual modo, Galán Rodríguez (1999) analisa os marcadores discursivos e no que diz respeito aos marcadores que marcam as relações de causalidade destaca *por causa de (que)*, *a causa de (que)*, *por razón de (que)*, *debido a (que)*, *gracias a (que)*, *por culpa de (que)*, *en vista de (que)*, *teniendo en cuenta (que)*, *una vez que*, *dado que*, *supuesto que* y *ahora (que)*. O mesmo autor identifica ainda relações causais introduzidas por outras conjunções, como *cuando*, *si*, *ahora que*, como ilustram os exemplos seguintes:

(83) Gracias a que se acabó tarde terminamos el trabajo.

(84) Gracias a tu ayuda estamos peor que al principio.

(85) Cuando yo lo digo, será por algo.

(86) Si ha estado allí, traerá noticias.

(87) Ahora que tengo dinero, podré hacer el viaje.

Galán Rodríguez (1999, pp. 3616-3620)

Nestes últimos casos, a causalidade depende de ligações semânticas entre dois enunciados e da alternância temporal, em que o primeiro enunciado exprime uma circunstância presente ou atual e o segundo indica uma consequência hipotética. Nestes contextos, estas conjunções perdem o seu valor semântico original e adquirem um valor causal atemporal, conforme discutido por Cunha & Silvano (2009) para o *quando*.

Galán Rodríguez (1999) mostra ainda que existem casos em que a causa e o efeito ocorrem simultaneamente, podendo a relação causal ser introduzida pela estrutura *al + infinitivo*, que, em certos contextos, é funcionalmente equivalente à

conjunção causal *como*. Contudo, enquanto *como* tende a apresentar a causa como objetiva ou necessária, a estrutura *al + infinitivo* remete para uma causalidade inferida, dependente da interpretação subjetiva do falante (Galán Rodríguez, 1999). Os exemplos seguintes ilustram estas leituras.

(88) Como no llamaste, nos fuimos

(89) Al no lamar tú, nos fuimos.

Galán Rodríguez (1999, p. 3620)

Para o português, Lopes & Carrilho (2013) distinguem entre estratégias de construção de coesão interoracional e interfrásica, as quais de incluem a coordenação e subordinação, que discutiremos no subcapítulo seguinte, e as da coesão construída com conectores que dão origem a unidades não frásica. Neste grupo, propõem seis categorias. As autoras identificam, em primeiro lugar, os marcadores reformuladores, cuja função consiste em retomar um segmento discursivo anterior para o clarificar, parafrasear ou sintetizar e alguns exemplos são como *ou seja, em resumo*. A segunda categoria corresponde aos marcadores especificativos, que têm como principal função a exemplificação ou particularização da informação previamente apresentada. Estes marcadores permitem restringir ou ilustrar o conteúdo do enunciado anterior, introduzindo exemplos ou casos específicos, como *por exemplo, como*. As autoras distinguem ainda os marcadores reforçativos, responsáveis pelo reforço da informação com valor aditivo, sendo exemplos *além disso* ou *de facto*. Outra categoria relevante é a dos marcadores estruturadores, cuja função principal é a ordenação da informação no discurso. Estes marcadores organizam a progressão textual, sinalizando sequências, mudanças de ponto ou fases do texto, como em *primeiro lugar, em seguida* ou *por fim*. Os marcadores consequenciais constituem uma categoria central na expressão das relações lógicas do discurso, sendo responsáveis pela indicação de consequência ou resultado. Estes marcadores estabelecem uma relação inferencial entre enunciados, apresentando um segmento como decorrente de outro, através de formas como *assim, desta forma, portanto*. Por fim, Lopes & Carrilho (2013) identificam os marcadores

contrastivos, cuja função é a expressão de contraste ou oposição entre segmentos discursivos. Estes marcadores introduzem um enunciado que contraria, limita ou relativiza a informação anterior, como acontece com *no entanto* e *contudo*.

Lopes & Carrilho (2020) apresentam uma proposta que tem alguns pontos de contacto com a de Lopes & Carrilho (2013), embora haja diferenças, nomeadamente em categorias como marcadores discursivos elaborativos, contrastivos e adversativos. Para além disso, contrariamente ao que acontece na proposta de (2013), Lopes & Carrilho (2020) consideram conjunções e locuções conjuncionais no grupo dos marcadores discursivos. Esta integração verifica-se na classe dos marcadores explicativos. É de referir que vários autores incluem no grupo de marcadores discursivos as conjunções ou locuções conjuncionais coordenativas e subordinativas, assumindo uma visão mais abrangente do que são marcadores discursivos, nomeadamente no quadro das relações discursivas. Como no subcapítulo seguinte, iremos apresentar as orações como estruturas que podem marcar as relações discursivas, não apresentaremos nesta secção uma revisão destes elementos linguísticos.

Centrando o foco nos marcadores com valor causal, que constituem o foco desta investigação, e voltando às propostas de Lopes e Castilho, enquanto a proposta de 2013 apenas estabelece a categoria dos marcadores consequenciais, a de 2020, para além desta, propõe ainda a classe dos conclusivos e explicativos. As autoras analisam discursos explicativos e conclusivos e consideram que estes envolvem “um raciocínio de natureza inferencial” e “funcionam em espelho, mantendo entre si uma relação de paráfrase” (Lopes & Carrilho, 2020, p. 2692). Nas construções conclusivas, o enunciado introduz uma conclusão ou consequência lógica inferida a partir de um enunciado anterior. Nas explicativas, os marcadores sinalizam que o enunciado subsequente deve ser interpretado como explicação ou justificação do que foi previamente afirmado, como acontece com *pois*, *porque*, *visto que*, *dado que*, *uma vez que*, *já que* ou *que*.

Outros autores, como Lobo (2003), debruçam-se sobre o estudo dos conetores com valor causal. Lobo (2003) mostra que as construções causativas e consequenciais podem ser marcadas por conectores como *que*, *pois*, *porque*, *visto que*, *dado que*, *já que*,

uma vez que, como, por + infinitivo, por causa de + infinitivo, visto + infinitivo e dado + infinitivo, muitos associados a construções de subordinação adverbial.

Mendes et al. (2023; 2024) apresentam um léxico de marcadores das relações discursivas em português europeu, ao aliarem uma fundamentação teórica sólida a recursos linguísticos anotados de forma sistemática. A criação do LDM-PT (*Lexicon of Discourse Markers for Portuguese*) (Mendes et al., 2024) representa um avanço relevante, ao disponibilizar um léxico extensivo de marcadores discursivos associados a relações do tipo PDTB, contemplando fenómenos de polissemia e ambiguidade categorial, frequentes em itens como *e, mas* ou *se*. Complementarmente, o estudo baseado no TED-MDB (Mendes et al., 2023) demonstra a relevância da articulação entre informação lexical e sintática na identificação automática de conectores discursivos, evidenciando que a posição, a estrutura frásica e o contexto discursivo são determinantes para a correta classificação funcional destes elementos.

3.4. Orações

De acordo com Raposo (2003, p. 306), “as noções de frase, enunciado e proposição encontram-se de tal maneira ligadas entre si que é praticamente impossível discutir cada uma delas sem mencionar as outras”. Uma frase pode ser entendida como:

uma sequência de palavras numa determinada ordem, que satisfaz as regras e os princípios gramaticais da língua a que pertence, e que descreve uma situação do mundo sobre o qual se fala ou remete para ela. As frases são elementos abstratos, inteiramente determinados pelo sistema gramatical; o seu estatuto é independente de qualquer realização oral ou escrita, ou seja, do seu uso concreto por falantes ou escritores, num momento e num lugar determinados. Assim, para dar um exemplo, a sequência de palavras os assaltantes do banco fugiram num carro, independentemente de ser dita ou escrita por alguém num momento particular, é uma frase, visto que descreve uma situação e satisfaz as regras e os princípios da gramática entre outros, p.e., a divisão em sintagma nominal seguido de sintagma verbal, a concordância em pessoa e número entre o sujeito e o verbo e a ordem linear artigo+ nome nas expressões nominais (Raposo, 2003, P. 306).

Relativamente ao enunciado, trata-se de uma unidade gramatical resultante de um ato de fala ou de escrita realizado por um falante ou escritor, contendo uma intenção comunicativa dirigida a um interlocutor ou a um grupo de interlocutores, numa determinada situação comunicativa. Segundo os mesmos autores, os enunciados,

enquanto produtos de um ato de fala ou de escrita, são elementos que podem ser percebidos (ainda que de modo efêmero, no caso da língua falada), encontrando-se inscritos num meio físico, de natureza sonora, no caso da fala, ou gráfica, no caso da escrita. Assim, sempre que uma frase é produzida num ato concreto de fala ou de escrita, num lugar e momento determinados, por um falante ou escritor específico, referindo uma situação particular de fuga de assaltantes de um banco num carro, constitui um enunciado (Raposo, 2003).

Já as proposições correspondem ao conteúdo expresso quando se produz um enunciado completo. Por outras palavras, “a frase enquanto unidade do sistema gramatical, independentemente da sua realização em enunciados concretos produzidos em situações particulares, dizemos que tem um conteúdo proposicional” (Raposo, 2003, p. 307). Os autores esclarecem ainda que “a noção de proposição se associa estreitamente com a noção de enunciado (de uma frase), ao passo que a noção de conteúdo proposicional se associa estreitamente com a noção de frase, enquanto unidade abstrata da gramática” (Raposo, 2003, p. 307).

Relativamente às orações, pode considerar-se que são enunciados que contêm verbos ou locuções verbais e que se estruturam em torno de um sujeito e de um predicado, ou apenas de um predicado. Neste sentido, uma oração é definida como uma “construção sintática, uma oração é uma sequência de palavras gramatical, que tem um conteúdo proposicional e como elemento nuclear um verbo” (Raposo, 2003, p. 314). É a partir do domínio proposicional verbal presente na frase que se identificam as orações, entendidas como unidades centrais frequentemente introduzidas por conjunções ou locuções conjuncionais, “vocábulos gramaticais que servem para relacionar duas orações ou dois termos semelhantes da mesma oração” (Cunha & Cintra, 1984, p. 575).

Deste modo, as conjunções permitem articular duas orações ou frases simples, dando origem a orações ou frases complexas. Tal como sublinha Raposo (2003, p. 315), “quando nenhuma das unidades que constitui uma oração é também uma oração, chamamos-lhe oração simples (...). Em contrapartida, quando uma oração contém outra oração que desempenha nela uma função gramatical, chamamos-lhe oração complexa”.

No grupo das orações, distinguem-se dois grandes grupos de orações complexas: as formadas por coordenação e as formadas por subordinação. Com o objetivo de distinguir estas duas estruturas, Lobo (2003) assinala que não existem critérios totalmente uniformes para diferenciar coordenação e subordinação. Ainda assim, a autora propõe uma classificação das orações introduzidas por diferentes conjunções com base em sete critérios (Lobo, 2003, p. 37):

- a. (in)existência de estrutura de encaixe
- b. (im)possibilidade de anteposição (ocorrência em posição inicial de frase)
- c. (im)possibilidade de extração do seu interior
- d. (im)possibilidade de coordenação
- e. (im)possibilidade de o conector ligar constituintes não frásicos
- f. (im)possibilidade de o conector ligar mais do que dois constituintes
- g. colocação dos clíticos (no caso das orações finitas)

Estes critérios são aplicados por Lobo (2003) à análise de cada tipo de oração.

Os resultados dessa análise encontram-se sintetizados no quadro seguinte.

Tabela 7 - Testes sintáticos para análise do comportamento de diferentes tipos de orações e conectores

	Estrutura de encaixe	anteposiç.	extracç.	coordenaç.	próclise	ligação de constit. não oracionais	ligação de + do que 2 elementos
copulativas (<i>e</i>)	*	*	*	*	*	√	√
disjuntivas (<i>ou</i>)	*	*	*	*	* ²⁴	√	√
adversativas (<i>mas</i>)	*	*	*	*	*	√	*
temporais	?	√	*	√	√	*	*
causais	?	√	*	√	√	*	*
finais	?	√	*	√	√	*	*
condicionais	?	√	*	√	√	*	*
concessivas	?	√	*	√	√	*	*
comparativas	√	*	*	?√	√	*	*
consecutivas	√	*	*	?*	√	*	*
relativas (<i>c/ ant. n.</i>)	√	*	*	√	√	*	*
completivas	√	√	√	√	√	*	*

Fonte: (Lobo, 2003, p. 39)

Segundo a autora, esta análise permite distinguir um conjunto de propriedades que delimitam claramente o grupo das orações subordinadas adverbiais (com a exceção das comparativas e consecutivas).

Orações com valor causal podem apresentar-se sob forma coordenada ou subordinada. Nem sempre é fácil distinguir as orações coordenadas explicativas das subordinadas causais, pelo que vários autores propuseram critérios para essa diferenciação.

De modo geral, define-se a coordenada explicativa como a oração que “exprime o motivo de se ter feito a declaração anterior” (Kury, 1985, p. 69, apud, Abreu et al., 2016) e a subordinada causal como aquela que “exprime uma circunstância causal, equivalente a um adjunto adverbial” (Abreu et al., 2016, p. 122). Normalmente, identifica-se o tipo de oração causal pelo conector que a introduz (conjunções ou locuções conjuncionais). Os exemplos retirados de Cunha & Cintra (1984, pp. 594-601) ilustram esta observação:

(90) Eh, camarada, espere um pouco que isto acaba-se já. (Fernando Namora)

(91) “Ceamos à lareira, que a noite estava fria. (Aquilino Ribeiro)

Ambas as frases apresentam a conjunção “que” introduzindo a segunda oração com valor causal. Contudo, (90) é classificada como oração coordenada explicativa, enquanto (91) é uma oração subordinada adverbial causal. A diferença torna-se clara ao analisarmos o conteúdo proposicional de cada construção. Em (90), a oração introduzida por “que” funciona como uma justificação para o ato de fala do interlocutor, isto é, explica por que o falante pede ao camarada que espere um pouco. Já em (91), a oração iniciada por “que” fornece o motivo da ação descrita na oração principal (cear à lareira). Um caso semelhante ocorre com o conector “pois”, que também pode introduzir orações explicativas. Comparem-se os exemplos de Cunha & Cintra (1984, pp. 582-594)

(92) “Tio Couto estava sombrio, pois aparecera um investigador da polícia perguntando por Gervásio. (Érico Veríssimo)

(93) “Um pouquinho só lhe bastava no momento, pois estava com fome. (Aníbal M. Machado

Ambas as construções são introduzidas por “pois”, mas há diferenças na interpretação. Em (92), a oração iniciada por “pois” apresenta uma causa factual que explica o estado de espírito de Tio Couto, ou seja, justifica a inferência do falante acerca do motivo da preocupação de Tio Couto (o facto de a polícia ter aparecido explica por que ele estava sombrio). Por isso, Cunha & Cintra (1984) classificam esta conjunção como subordinativa causal. Já em (93), a oração introduzida por “pois” serve de justificação para a afirmação do falante de que “um pouquinho só lhe bastava”: o falante crê que, por ele estar com fome, um pequeno alimento seria suficiente naquele momento. Neste caso, “pois” é classificado como conjunção coordenativa explicativa.

Esses exemplos ilustram a complexidade das relações causais e de justificação na língua. De facto, Dik et al. (1990) e Dik (1997a) já haviam salientado que, sob o rótulo genérico de “relação de causa”, existem relações distintas – como causa, razão e explicação – que não devem ser confundidas. Segundo esses autores, *“causa indica que o evento expresso na oração dependente desencadeia a ocorrência do evento expresso na oração principal, sem que haja qualquer envolvimento intencional por parte de um agente”* (Dik et al., 1990; Dik, 1997a, *apud* Abreu et al., 2016, p. 123). Por outras palavras, a causa propriamente dita refere-se a uma relação causal direta entre dois eventos, excluindo motivações ou intenções humanas.

A dificuldade de distinção entre as orações coordenadas explicativas e as subordinadas causais, pelo que vários autores propuseram critérios para essa diferenciação. Kury (1985) (*apud* Abreu et al. (2016, p. 122), por exemplo, propõe cinco critérios específicos para distinguir uma oração subordinada adverbial causal de uma coordenada explicativa. Esses critérios são:

- 1) a subordinada vale por um adjunto adverbial, o que não ocorre com a coordenada, que é sintaticamente independente; 2) a oração explicativa, por ser independente, admite pausa forte, que pode ser indicada, na escrita, por dois pontos ou ponto e vírgula; 3) o conectivo, nas explicativas, pode ser omitido sem qualquer prejuízo da clareza; 4) na maioria das vezes, a oração que antecede uma explicativa tem o verbo no imperativo; 5) em sua maioria, as orações causais iniciadas com “que”, “pois”,

“porque” podem ser substituídas por equivalentes com os conectivos “como”, “uma vez que” e análogos, o que não é possível com as explicativas (Kury, 1985, *apud*, Pezatti, Novaes & Marques, 2016, p. 122).

Lobo (2003) elenca quatro critérios principais, apresentados na seguinte tabela.

Tabela 8 - Critérios de distinção do estatuto mais coordenativo ou subordinativo dos conetores causais/explicativos

conjunções de coordenação e tipos de orações causais/conectores causais	testes sintáticos			
	possibilid. de ocupar posição inicial	possibilid. de ser coorden.	possibilid. de ligar constit. não frásicos	posição dos clíticos: próclise (vs. ênclise)
<i>e</i>	*	*	√	*
<i>ou</i>	* ^a	*	√	* ³²
<i>mas</i>	*	*	√	*
<i>que</i>	*	*?	*	*?
<i>pois</i>	*	*	*	*
<i>porque</i> (explicativo)	*	*?	*	*?
<i>porque</i>	√	√	*	√?
<i>visto que</i>	√	√ ³³	*	√
<i>dado que</i>	√	√	*	√
<i>já que</i>	√	√	*	√
<i>uma vez que</i>	√	√	*	√
<i>como</i>	√	√	*	√
<i>por</i> + infinitivo	√	√	*	-
<i>por causa de</i> + inf ³⁴	√	√	*	-
<i>visto</i> + inf.	√	√	*	-
<i>dado</i> + inf.	√	√	*	-

Fonte: (Lobo, 2003, p. 51)

Os testes sintáticos apresentados na Tabela mostram que *pois*, *que* e *porque* (explicativo) têm comportamentos diferentes dos restantes conectores causais investigados, partilhando mais características com os conectores coordenativos (Lobo, 2003). A análise de Lobo (2003) permite ainda concluir que as orações coordenadas explicativas, por sua vez, têm propriedades que as afastam das orações coordenadas, mas que também não validam a sua integração no grupo das subordinadas.

Outra distinção importante na análise semântica das construções causais é apontada por Lopes (2004), que propõe duas oposições semânticas complementares: (a) causa explicativa vs. causa não-explicativa e (b) causa *de re* vs. causa *de dicto*. A causa explicativa refere-se às construções coordenadas explicativas tradicionais (ligadas ao ato de fala e muitas vezes *de dicto*), enquanto causa não-explicativa diz respeito às subordinadas causais de motivo/razão (ligadas a fatos do mundo, frequentemente *de re*). Ambas as categorias, no entanto, podem manifestar causas do tipo *de re* (facto objetivo) ou *de dicto* (inferência ou avaliação do falante). Vejamos os exemplos adaptados de Lopes (2004, pp. 37-38) que ilustram essas combinações:

- (94) A janela está estragada, pois a dobradiça partiu. – (Causa explicativa com interpretação de *re*: a segunda oração enuncia um facto real, a dobradiça partiu, que explica objetivamente o estado da janela.)
- (95) A janela está estragada, pois não a consigo abrir. – (Causa explicativa com interpretação de *dicto*: aqui a causa expressa é subjetiva, “não consigo abri-la”, indicando a inferência do falante para concluir que a janela está estragada, em vez de um facto físico comprovado.)
- (96) Foi idiota por não aceitar a proposta. – (Causa não-explicativa com valor de *re*: “por não aceitar a proposta” é apresentado como um facto concreto que qualifica a ação, não aceitar a proposta foi um ato idiota em si mesmo, objetivamente considerado.)
- (97) Acho-o idiota por não aceitar a proposta. – (Causa não-explicativa com valor de *dicto*: nesta frase, “por não aceitar a proposta” reflete a razão subjetiva na perspetiva do falante para julgá-lo idiota; é a opinião do falante baseada naquele facto, e não uma idiotice intrínseca ao ato.)

No caso das orações subordinadas causais de razão (causa não-explicativa), na análise de Silvano (2010), esta classificação de Lopes (2004) coincide em grande medida com a tipologia bipartida tradicional, pois tais orações tratam de causas de factos (*de re* quando objetivas, *de dicto* quando inferidas). Já no tocante às orações coordenadas explicativas, a questão torna-se mais controversa. Alguns autores, como Lobo (2003),

defendem que as coordenadas explicativas expressam apenas causas do tipo *de dicto* (isto é, justificativas do ponto de vista do falante, não causas objetivas em si). Outros, como a própria Lopes (2004), argumentam que mesmo construções coordenadas explicativas podem veicular causas *de re* (isto é, factos objetivos que explicam o enunciado anterior). De qualquer modo, não nos alongaremos nessa controvérsia, visto que o enfoque do presente trabalho recai apenas sobre orações em que a causa é interpretada como *de re* (isto é, causas objetivas, pertencentes ao domínio dos estados de coisas).

Além das construções tradicionais com conectores causais explícitos, as relações de causalidade também podem ser expressas de forma implícita ou por meio de construções reduzidas. Na sua investigação, Nuss (2021) observa que as relações causais, englobando tanto a *causa* propriamente dita quanto a *explicação*, se manifestam através de construções prototípicas variadas. Essas construções incluem as orações causais desenvolvidas (sejam coordenadas ou subordinadas com conjunção) e as orações reduzidas de infinitivo, gerúndio ou particípio (conforme já atestado em Cunha & Cintra, 1984; Lobo, 2003, entre outros).

Lobo (2003) destaca que as orações subordinadas causais podem aparecer em formas não canónicas, e que certos conectores causais já trazem embutido um traço semântico-discursivo especial. Por exemplo, conectores como “*visto que*”, “*dado que*”, “*uma vez que*” e “*já que*” carregam no léxico um traço [+pressuposicional] (Lobo, 2003). Isso significa que, ao usar “*já que*”, por exemplo, o falante assume como pressuposto verdadeiro o conteúdo da oração introduzida por esse conector, estabelecendo uma causa que não se apresenta como informação nova, mas sim como conhecimento compartilhado ou facto dado como sabido. Por outro lado, construções causais sem conector explícito são consideradas *subespecificadas* em termos semântico-discursivos (Lobo, 2003). Nestas, o valor exato da relação causal depende em grande medida da posição da oração na frase e da interpretação pragmática do contexto. Por exemplo, uma oração causativa anteposta sem conjunção pode ser entendida como causa ou condição, conforme a entoação e a situação comunicativa.

Esta questão das causais implícitas foi aprofundada por Meyer (2000), que discute a frequente interpretação causal de sequências de orações justapostas. Meyer salienta que, na literatura, a causalidade é um tema controverso, mas observa que muitas vezes duas orações ligadas sem conector claro são interpretadas como tendo relação causal, seja através de justaposição simples, de coordenação com *e*, ou mesmo em construções de particípio, como ilustram os seguintes exemplos adaptados de Meyer (2000):

(98) “People badly want a demonstration, a dramatization of justice. They need to have defined and reinforced for them what is right and what is wrong.”

(99) “The accusatorial system is more adaptable than the inquisitorial one, allowing for practical changes in treatment methods to be introduced easily during the sentence.”

(100) “The participants remain anonymous and there is no one upon whom the authorities can pin the offence.”

Nos exemplos acima, apesar de não haver conectores causais explícitos, a segunda oração de cada par funciona como uma explicação ou justificação da primeira. Em (98), por exemplo, a frase “*They need to have defined and reinforced for them what is right and what is wrong*” explica por que as pessoas desejam “*uma dramatização da justiça*”, ou seja, fornece a razão motivadora dessa necessidade. O mesmo padrão explicativo pode ser observado em (99) e (100), nos quais a frase após o ponto final é interpretada como elucidando ou fundamentando a afirmação precedente. No entanto, Meyer ressalva que, em certos contextos, uma segunda oração justaposta pode também ser entendida como consequência da primeira, em vez de sua causa explicativa, como no seguinte exemplo.

(101) “In all the countries concerned the population is growing quickly. All will be faced sooner or later with the specter of overpopulation.”

Neste caso, a segunda frase (*"All will be faced... overpopulation"*) pode ser lida como uma consequência previsível da situação descrita na primeira frase (o rápido crescimento populacional).

Em suma, a análise das orações causais em português requer a consideração de múltiplos fatores, sintáticos, semânticos, pragmáticos e mesmo prosódicos, para se determinar com rigor que tipo de relação está expresso. A distinção entre orações coordenadas explicativas e subordinadas causais não se resume à presença ou ausência de determinada conjunção, mas envolve compreender o papel da oração no discurso (se justifica um ato de fala ou relata um motivo factual), o estatuto das informações veiculadas (pressupostas, inferidas ou novas), bem como a lógica interna do enunciado (causa objetiva vs. explicação subjetiva). Podemos reconhecer que "causa" não é um conceito monolítico, mas abrange uma família de relações, desde causas físicas diretas até razões cognitivas e justificativas discursivas, que a língua codifica de formas diversas.

3.5. Conclusão

Neste capítulo foi nosso objetivo refletir sobre alguns elementos e estruturas linguísticas que têm valor causal e que podem contribuir para a inferência das relações causais. Focamos a nossa análise em elementos lexicais como os verbos causativos, nomes e advérbios com informação de natureza causativa no seu significado, assim como preposições que podem ter interpretação causal e, por fim, em orações com valor causal.

Esta reflexão permitiu-nos concluir que há uma diversidade de elementos e estruturas linguísticas que podem efetivamente funcionar como marcas da relação Causa. Contudo, como na nossa análise dos dados iremos considerar, seguindo a ISO 24617-8, relações discursivas entre duas ou mais situações (sejam eventos ou estados), haverá alguns casos de verbos, nomes e advérbios causativos que não consideraremos como sinais da relação Causa.

A revisão feita revela o papel fundamental que o léxico tem na construção do significado. Contudo, como foi discutido neste capítulo e também no anterior, a interpretação das unidades lexicais não pode ser explicada exclusivamente através de informação lexical isolada. Pelo contrário, a compreensão do significado resulta da interação de conhecimento lexical, conhecimento semântico, conhecimento

enciclopédico e estrutura discursiva.

4. As categorias de Tempo e Aspeto

O Tempo e o Aspeto constituem duas dimensões fundamentais no domínio da Semântica, sendo geralmente considerados conceitos complexos, na medida em que a sua realização linguística é condicionada por múltiplos fatores. Nesta perspetiva, torna-se essencial compreender as suas definições e propriedades, de modo a alcançar uma interpretação rigorosa das relações temporais que se estabelecem entre os eventos no discurso.

Para uma caracterização adequada dessas relações temporais, é imprescindível considerar a interação entre Tempo e Aspeto. Com efeito, como refere Oliveira (2003, p. 509), a “localização temporal e a estrutura aspetual dos enunciados influenciam-se mutuamente; em particular, é frequente o valor temporal expresso por um determinado tempo verbal depender dos valores aspetuais veiculados na frase” (Oliveira, 2003, p. 509). Esta interdependência evidencia a necessidade de uma abordagem integrada destas duas categorias gramaticais.

Uma vez que o objetivo deste trabalho consiste em analisar a organização temporal das situações que configuram relações causais no discurso, revela-se pertinente proceder a uma revisão das principais propostas teóricas centradas no estudo do Tempo e do Aspeto. Tal enquadramento permitirá sustentar, de forma consistente, a análise dos dados empíricos.

Assim, o presente capítulo tem como objetivo apresentar as noções de Tempo e de Aspeto, com vista à construção de uma base teórica sólida que possibilite, numa fase posterior, a exploração do contributo das relações temporais para a análise das relações causais no discurso. Nesse sentido, serão inicialmente apresentadas as noções gerais destas duas categorias e as suas principais propriedades. Seguidamente, serão discutidas algumas propostas teóricas relevantes para a sua caracterização, nomeadamente, no que diz respeito ao Tempo, as de Reichenbach (1947), Kamp & Reyle (1993) e Oliveira (2003; 2013), e, no domínio do Aspeto, as de Vendler (1967), Moens & Steedman (1988) e Cunha (2013).

4.1. Noção de Tempo e Aspeto

As categorias Tempo e Aspeto constituem dimensões fundamentais para a organização e a interpretação dos eventos no eixo temporal das línguas naturais. Enquanto o Tempo se relaciona com a localização das situações relativamente a um determinado intervalo de tempo, o Aspeto diz respeito à forma como a estrutura interna dessas situações é concetualizada e apresentada linguisticamente.

De acordo com Oliveira (2003, p. 129), “a categoria Tempo serve para localizar as situações (eventos ou estados) expressas na língua em diferentes tipos de enunciados”. Essa localização é efetuada sempre em relação ao tempo da enunciação ou a um tempo de referência, sendo codificada por diversos recursos linguísticos, nomeadamente os tempos verbais, os verbos auxiliares e semiauxiliares, os adverbiais temporais e determinadas orações com valor temporal.

No que diz respeito ao Aspeto, este “fornece informações sobre a forma como é perspectivada ou focalizada a estrutura temporal interna de uma situação descrita pela frase, em particular, pela sua predicação” (Oliveira, 2003, p. 129). O Aspeto não determina quando uma situação ocorre, mas antes como essa situação se distribui no tempo, atendendo a propriedades como dinamismo, telicidade, duração.

Segundo Cunha (2004), as categorias de Tempo e de Aspeto podem assentar numa mesma base concetual, estruturada em torno de momentos e intervalos temporais, permitindo responder, respetivamente, às questões de quando ocorre uma situação e de como essa situação se desenvolve ao longo do tempo. Esta perspectiva encontra-se já em Bennett (1988), para quem os valores temporais e aspetuais de um evento podem ser descritos exclusivamente com base em noções temporais, como instantes e intervalos. Nesta abordagem, o Aspeto é concebido como uma dimensão derivável de conceitos temporais, ainda que com uma função interpretativa distinta.

A complementaridade, bem como o papel diferenciado, que Tempo e Aspeto desempenham na interpretação das predicções manifesta-se de forma particularmente clara nos tempos verbais. De acordo com Cunha (2004), os tempos gramaticais não se limitam a situar os eventos na linha temporal, mas veiculam também valores aspetuais. O autor ilustra esta articulação através do uso do Presente e do

Pretérito Imperfeito do indicativo. Do ponto de vista temporal, o primeiro localiza a situação no presente, enquanto o segundo a situa no passado. No entanto, ambos expressam o mesmo valor aspetual, como se observa nos exemplos seguintes:

(102) A Ana fuma.

(Oliveira, 2013, p. 515)

(103) A Ana fumava.

(Oliveira, 2013, p. 521)

Em ambos os casos, os tempos verbais veiculam um valor aspetual habitual. No exemplo (102), a frase indica que Maria tem o hábito de fumar, ou seja, descreve uma situação reiterada que se mantém válida no momento da enunciação. No exemplo (103), exprime-se que Maria tinha esse hábito num intervalo temporal anterior ao presente, caracterizando igualmente uma situação contínua ou reiterada no passado.

Os exemplos seguintes ilustram, uma vez mais, a importância de considerar simultaneamente as categorias de Tempo e de Aspeto na interpretação das predicções:

(104) A Ana casou em 1969. (localização temporal)

(105) A Ana mudou-se para Lisboa durante a década de 60. (delimitação aspetual)

(Móia & Alves, 2003, p. 564)

Nestes exemplos, os adverbiais temporais desempenham funções distintas. Em (104), o sintagma *em 1969* funciona como um localizador temporal, situando o evento num intervalo específico da linha do tempo. Em (105), pelo contrário, a expressão *durante a década de 60* atua como um delimitador aspetual, especificando a extensão interna da situação e contribuindo para a determinação do seu valor aspetual. Assim, embora ambos os adverbiais façam referência ao tempo, apenas no primeiro caso se trata de uma localização temporal externa, enquanto no segundo se evidencia uma delimitação da duração da situação.

Apesar da estreita relação entre Tempo e Aspeto, vários estudos sublinham a necessidade de distingui-los conceitualmente. Oliveira (2003), por exemplo, define o Tempo como “uma ordenação linear de unidades temporais atômicas (instantes) ou densas (intervalos) que se podem suceder ou sobrepor” (p. 129). Esta conceção permite organizar os eventos numa linha cronológica, estabelecendo relações de sucessão, quando uma situação ocorre depois de outra, ou de sobreposição, quando uma situação decorre no interior do intervalo de outra.

O Aspeto, por sua vez, não incide sobre a posição temporal externa dos eventos, mas sobre a sua estrutura interna, analisando a forma como as situações se desenvolvem no tempo a partir do seu interior. Uma outra diferença relevante entre estas categorias reside no seu estatuto relacional. Os tempos verbais ou linguísticos são elementos de natureza relacional (Reichenbach, 1947), uma vez que estabelecem relações dêiticas ou anafóricas com o tempo da enunciação (ponto de fala) ou com um tempo previamente introduzido no discurso (ponto de referência). Oliveira (2003) defende igualmente que esta ancoragem relacional é característica do Tempo, enquanto o Aspeto não exige, por si só, esse tipo de referência externa.

4.2. Algumas propostas teóricas de Tempo

Neste subcapítulo, apresentar-se-á uma revisão de algumas das propostas mais relevantes para a investigação de questões de Tempo.

4.2.1. Reichenbach (1947)

Tendo em consideração que os tempos verbais veiculam estruturas temporais complexas, Reichenbach (1947) propõe uma análise a partir de três pontos fundamentais: *the point of speech* (ponto de fala), *the point of the event* (ponto de evento) e *the point of reference* (ponto de referência). O ponto da fala corresponde ao momento da enunciação do ato de fala, o ponto do evento refere-se ao tempo em que ocorre o acontecimento descrito pela frase e o ponto de referência designa um ponto intermédio a partir do qual o acontecimento pode ser situado.

Com base nas combinações possíveis entre estes três pontos temporais, Reichenbach (1947) recorre a uma abordagem bidimensional do tempo verbal para

definir nove tempos verbais distintos na língua inglesa. No âmbito desta teoria temporal bidimensional, o autor introduz uma nova terminologia centrada, por um lado, nas relações entre o ponto de referência (R) e o ponto da fala (S), que podem ser de passado ($R < S$), presente ($R = S$) ou futuro ($R > S$), e, por outro, nas relações entre o ponto do evento (E) e o ponto de referência (R), que podem ser de anterioridade ($E < R$), simultaneidade ($E = R$) ou posterioridade ($E > R$). Estas relações encontram-se sintetizadas na tabela seguinte.

Tabela 9 - Caracterização dos tempos verbais segundo Reichenbach (1947)

Structure	New Name	Traditional Name
E-R-S	Anterior Past	Past Perfect
E,R-S	Simple Past	Simple Past
R-E-S	Posterior Past	
R-S,E		
R-S-E		
E-S,R	Anterior Present	Present Perfect
S,R,E	Simple Present	Present
S,R-E	Posterior Present	Simple Future
S-R-E	Anterior Future	Future Perfect
S,E-R		
E-S-R		
S-R,E	Simple Future	Simple Future
S-E-R	Posterior Future	

Fonte: adaptada de Reichenbach (1947, p. 297)

Para representar a coesão temporal, no âmbito da sequência de tempos, o autor propõe dois princípios: o princípio da permanência do ponto de referência e o princípio do uso posicional do ponto de referência. Quanto ao primeiro princípio, este é formulado nos seguintes termos: “although the events referred to in the clauses may occupy different time points, the reference point should be the same for all clauses—a principle which (...) demands the permanence of the reference point” (Reichenbach, 1947, p. 74). Tal como se pode observar no exemplo seguinte:

(106) I had mailed the letter when John came and told me the news.

1st clause: E1—R1 —S

2nd clause: R2, E2—S

3rd clause: R3, E3—S

A partir deste exemplo, identifica-se uma sequência de eventos expressos por tempos verbais distintos, nomeadamente o evento da chegada de John e o evento de contar a notícia. Embora estes eventos ocorram em momentos diferentes, o ponto de referência mantém-se comum aos dois, garantindo a coerência temporal entre as orações.

Relativamente ao princípio do uso posicional do ponto de referência, este é aplicado “if the time relation of the reference points compared is not identity, but time sequence, i.e., if one is said to be before the other, the rule of the permanence of the reference point can thus no longer be maintained” (Reichenbach, 1947, p. 75). O exemplo seguinte ilustra um desses casos.

(107) ‘he telephoned before he came’

(Reichenbach, 1947, p. 75)

O mesmo princípio do uso posicional do ponto de referência permite ainda inferir explicitamente relações temporais de anterioridade, simultaneidade ou posterioridade em orações subordinadas temporais introduzidas por *quando*, *antes* ou *depois*.

A proposta de Reichenbach (1947) constitui um marco incontornável no estudo da expressão de Tempo nas línguas naturais, ao introduzir um modelo sistemático para a representação das relações temporais entre o momento da enunciação, o tempo do evento e um ponto de referência. Apesar do seu impacto duradouro, esta abordagem revela algumas limitações, em particular no recurso a um único ponto de referência, o que se mostra insuficiente para captar a complexidade das relações temporais envolvidas em enunciados mais elaborados, como aqueles que incluem sequências de tempo. Foi precisamente para ultrapassar essas limitações que Kamp & Reyle (1993) desenvolveram uma proposta que reformula e expande a noção de ponto de referência, permitindo uma representação mais adequada e completa das relações temporais, proposta que será desenvolvida na secção seguinte.

4.2.2. Kamp & Reyle (1993)

A teoria de Kamp & Reyle (1993) insere-se no domínio da Semântica e da lógica da linguagem natural e é designada por *Teoria da Representação do Discurso* (*Discourse Representation Theory* – DRT). Esta abordagem tem como objetivo dar resposta a problemas de interpretação da anáfora nominal e temporal, viabilizando a representação formal do contributo do Tempo e Aspecto para o processo de interpretação do discurso.

No que diz respeito à categoria de Tempo, os autores partem da abordagem de Reichenbach (1947), mas propõem o desdobramento do ponto de referência em reference point (Rpt) (ponto de referência) e o temporal perspective point (TPpt) (ponto de perspectiva temporal, ambos utilizados na interpretação temporal de frases isoladas ou de sequências de frases no discurso. O Ponto de Referência Temporal (Rpt) corresponde ao “type of reference time which accounts for narrative progression” (Kamp & Reyle, 1993, pp. 594-595), isto é, o parâmetro através do qual os eventos descritos se organizam uns em relação aos outros, seguindo uma sequência temporal lógica que assegura a progressão narrativa. O Ponto de Perspetiva Temporal (TPpt), por sua vez, é definido como o “time from which the described eventuality is seen” (Kamp & Reyle, 1993, p. 595). Para uma melhor compreensão, considere-se o exemplo seguinte:

(108) Fred arrived at 10. He had got up at 5; he had taken a long shower, had got dressed and had eaten a leisurely breakfast. He had left the house at 6:30.

Neste exemplo, observa-se claramente o funcionamento dos dois pontos, Rpt e TPpt. Por um lado, o pretérito mais-que-perfeito (Past Perfect) é utilizado para indicar que os eventos descritos (acordar, tomar banho, vestir-se, tomar o pequeno-almoço e sair de casa) ocorreram antes do Ponto de Perspetiva Temporal (TPpt), a partir do qual esses eventos são perspectivados, que corresponde ao momento da chegada de Fred. Ao longo da progressão narrativa, o TPpt mantém-se estável. Por outro lado, o Rpt, ou Ponto de Referência Temporal responsável pela progressão narrativa, altera-se ao longo do discurso, estruturando a sequência temporal dos eventos de forma clara e lógica.

Cada eventualidade passa a funcionar como ponto de referência para o evento seguinte, seguindo o eixo cronológico da narrativa e assegurando a coerência temporal entre os acontecimentos.

No que diz respeito à caracterização dos tempos verbais, os autores caracterizam recorrem às relações definidas segundo os princípios seguintes:

- TPpt (Ponto de Perspetiva Temporal) ao momento de enunciação (now)
- +PAST: O TPpt se localiza antes do momento de enunciação
- PAST: O TPpt sobrepõe-se ou coincide com o momento de enunciação
- Tempo de localização da eventualidade ao TPpt (Ponto de Perspetiva Temporal)
- O evento descrito ocorre antes, simultâneo, depois ao TPpt.

Para além disso, Kamp & Reyle (1993) caracterizam os tempos verbais quanto às suas propriedades aspetuais recorrendo ao valor estativo ou eventivo (+/-STAT) e ao valor de estado resultativo (+/-PERF).

A tabela seguinte sistematiza a caracterização dos tempos verbais de acordo com a proposta de Kamp & Reyle (1993).

Tabela 10 -Caracterização dos tempos verbais segundo Kamp & Reyle (1993)

Tempo Verbal	Relação entre TPpt e momento de enunciação (PAST+/-)	Relação entre tempo de localização da eventualidade descrita e TPpt	Eventivo (+/-STAT)	Estado Resultativo (+/-PERF)
Presente	- PAST	Pres	+STAT	- PERF
Futuro	- PAST	Fut	+/-STAT	- PERF
Pretérito Perfeito	- PAST	Past Pres	+/-STAT + STAT	- PERF - PERF
Futuro Passado	+ PAST	Fut	+/-STAT	- PERF
Present perfect	- PAST	Pres	+ STAT	+ PERF
Futuro Perfeito	+ PAST	Fut	+ STAT	+ PERF
Pretérito mais que perfeito	+ PAST	Past	+/-STAT	- PERF
	- PAST /+ PAST	Past Pres	+ STAT + STAT	+ PERF + PERF

Passado Futuro Perfeito	+ PAST	Fut	+ STAT	+ PERF
--	---------------	------------	---------------	---------------

Fonte: Silvano (2002, p. 47), adaptado de Kamp & Reyle (1993, p. 601)

4.2.3. Oliveira (2003; 2013)

Segundo Oliveira (2003), a categoria linguística do tempo pode ser representada por uma linha cronológica orientada do passado para o futuro ao longo do eixo temporal, estruturando-se em três dimensões gramaticais fundamentais: passado, presente e futuro. A organização temporal na linguagem dinamiza-se a partir do ponto da fala ou momento de enunciação, permitindo estabelecer relações temporais de anterioridade, simultaneidade ou posterioridade.

Uma vez que o Tempo é uma categoria relacional, ancorado no momento da enunciação ou num ponto de referência dado pelo contexto linguístico, as relações temporais estabelecidas podem ser dêiticas e anafóricas (Oliveira, 2003). No primeiro caso, a localização temporal é feita diretamente ao momento em que o enunciado é produzido. No segundo, a relação é determinada linguisticamente por um valor temporal previamente estabelecido, que funciona como ponto de referência para a situação principal. Estas duas situações podem ser observadas nos exemplos seguintes:

(109) Amanhã, A Maria está em casa.

(Oliveira, 2013, p. 516)

(110) A Maria chegou enquanto todos estavam a jantar.

(Oliveira, 2013, p. 513)

No primeiro exemplo, o advérbio *amanhã* refere-se a uma situação que ocorrerá num tempo futuro, mais concretamente no dia seguinte ao momento da enunciação. Trata-se, assim, de uma relação dêitica. No segundo exemplo, estabelece-se uma relação temporal entre os eventos *sair* e *parar*, sendo a situação descrita pela oração subordinada temporal o ponto de referência relativamente ao qual se localiza a situação da oração principal, o que caracteriza uma relação anafórica.

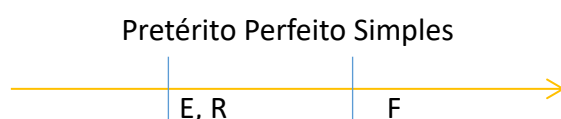
Tal como outros autores, Oliveira (2003) apresenta uma concepção temporal assente numa sequência de momentos organizados de forma linear e ordenada em intervalos. As relações temporais entre esses intervalos assumem três configurações principais: (i) o intervalo A sobrepõe-se ao intervalo B; (ii) o intervalo A está contido no intervalo B; ou (iii) o intervalo A precede o intervalo B. As duas primeiras configurações correspondem a valores temporais de simultaneidade, que podem manifestar-se como sobreposição parcial, total ou inclusão, enquanto a última expressa uma relação de anterioridade ou posterioridade. Para sistematizar estas interações temporais, e com base na perspetiva de Reichenbach (1947), Oliveira (2003, p. 131) identifica três momentos essenciais: o ponto da fala (F), que coincide com o momento da enunciação; o ponto do evento (E), que corresponde ao tempo do acontecimento descrito pela frase; e o ponto de referência (R), que funciona como um ponto intermediário a partir do qual se situa o evento ou estado descrito.

Considere-se o exemplo seguinte, no qual se recorre a estes três pontos para analisar as relações temporais, descrevendo-se uma situação que ocorre antes do momento da enunciação:

(111) A Maria escreveu a carta.

(Oliveira, 2013, p. 517)

Neste exemplo, a ação de escrever a carta localizou-se num tempo passado, pelo que o ponto do evento se localiza antes do ponto da fala, que corresponde ao momento presente. As relações temporais do exemplo anterior podem ser representadas pelo diagrama seguinte.



Para além do Pretérito Perfeito, Oliveira (2003; 2013) apresenta uma caracterização de alguns tempos verbais em português europeu, nomeadamente: (i) o Presente, (ii) o Pretérito Perfeito Composto, (iii) o Pretérito Imperfeito, (iv) o Pretérito

Mais-que-Perfeito, simples e composto, e (vi) o Futuro, simples e composto. Cada um destes tempos apresenta propriedades temporais e aspetuais específicas, que iremos expor brevemente no que se segue com base em Oliveira (2003; 2013).

O Presente do indicativo é, do ponto de vista teórico, o tempo deítico por excelência, uma vez que estabelece uma relação de sobreposição entre o tempo da situação e o momento da enunciação. No entanto, em português, essa leitura estritamente deítica é pouco frequente na ausência de adverbiais temporais ou de contextos específicos, como os relatos em direto ou as frases performativas (por exemplo, *Prometo que vou fazer mais exercício*). A leitura de Presente semântico (Oliveira, 2013) só ocorre quando a situação é estativa (*O Ricardo mora em Braga*). Com eventos, para exprimir de forma mais clara a simultaneidade entre a eventualidade e o momento da fala, recorre-se frequentemente à construção progressiva com o auxiliar estar (*O Pedro está a abrir a porta*). Para além disso, o Presente pode assumir valores não deíticos, como o Presente histórico (*Em 476, os bárbaros incendiam Roma*), leituras de futuro próximo (*Amanhã viajo para Macau*) ou interpretações instrucionais e modais, comuns em descrições de itinerários e receitas. Aspetualmente, a leitura preferencial do Presente, à exceção dos predicados estativos, é a habitual ou frequentativa, como em *O João corre* ou *O João corre a maratona todos os anos*. Oliveira (2013) mostra que, tal como acontece com o Pretérito Imperfeito, o Presente é um tempo aspetualmente imperfetivo, pois representa a situações no seu decurso.

O Pretérito Perfeito Simples localiza tipicamente uma situação anterior ao momento da enunciação, apresentando-a como terminada. Do ponto de vista aspetual, o Pretérito Perfeito mantém as categorias básicas das predicções, impondo-lhes apenas o traço de terminadas.

O Pretérito Imperfeito caracteriza-se por uma forte contribuição aspetual. Embora possa veicular informação de passado, este tempo verbal perspetiva frequentemente as situações como não delimitadas, em curso ou habituais. Em construções com outra situação pontual no passado, estabelece-se uma relação de sobreposição temporal (*O Pedro escrevia uma carta quando a Rita telefonou*), sendo a eventualidade no Imperfeito interpretada como incompleta. O Imperfeito favorece

ainda leituras habituais (*O Pedro jogava futebol aos domingos*) e pode ocorrer em contextos em que o ponto de referência não é passado, mas fornecido por advérbios como *agora* ou *logo*. Aspetualmente, este tempo tende a transformar eventos télicos em predicções atélicas ou estativas.

O Pretérito Perfeito Composto não expressa perfectividade, mas antes uma duração que se inicia no passado e se prolonga até ao presente. Dependendo do tipo aspetual do predicado, esta duração pode dar origem a leituras iterativas (*O Pedro tem nadado todos os dias*). Com estados, não há iteratividade, mantendo-se apenas a continuidade temporal (*O Pedro tem estado doente*). A interpretação iterativa deste tempo depende, em regra, da ancoragem no momento da enunciação e do modo indicativo.

O Futuro simples raramente exprime apenas posterioridade temporal em relação ao momento da enunciação, sendo frequentemente associado a valores modais, como possibilidade ou inferência (*Neste momento, os meus estarão a visitar a minha tia*). Em português, a posterioridade é mais frequentemente expressa pelo Presente ou por construções com o auxiliar *ir* (*A Ana vai visitar a mãe*). O Futuro composto, quando tem leitura temporal, exprime anterioridade relativamente a um ponto de referência futuro (*Quando chegares a casa, já o João terá partido*), podendo igualmente assumir leituras modais.

Por fim, o Pretérito mais-que-perfeito, simples ou composto, codifica a anterioridade de uma situação relativamente a outra situada no passado. Com predicções culminadas, permite estabelecer um estado conseqüente a partir do qual se localiza a situação seguinte (*A Maria tinha espirrado quando o João chegou*). Em certos contextos, o Pretérito mais-que-perfeito composto pode ainda atenuar afirmações ou pedidos, aproximando-se pragmaticamente do Imperfeito.

Em síntese, os tempos verbais do indicativo em português revelam uma articulação estreita entre valores temporais e aspetuais, sendo responsáveis não apenas pela localização cronológica das situações, mas também pela forma como estas são concetualizadas e integradas na progressão temporal do discurso.

4.3. Algumas propostas teóricas de Aspeto

O Aspeto, inicialmente associado sobretudo às propriedades dos verbos, revelou-se um fenómeno que interage com diversos elementos linguísticos e estruturais da predicação. Em muitas línguas, a informação aspetual manifesta-se através de afixos com valor temporal, construções com auxiliares e semiauxiliares, advérbios e até sintagmas nominais.

Neste contexto, surgiram várias abordagens teóricas que procuram identificar regularidades aspetuais nas situações e proceder à sua categorização em classes aspetuais. Diversos estudos contribuíram para este domínio, nomeadamente os de Vendler (1967), Dowty (1979), Moens (1987) e Moens & Steedman (1988). Neste trabalho iremos focar-nos nas propostas de Vendler (1967), Moens & Steedman (1988) e Cunha (2013), analisando os seus contributos para a compreensão de questões aspetuais.

4.3.1. Vendler (1967)

Vendler (1967) argumenta que os verbos não se limitam a contribuir para a localização temporal das situações por meio dos tempos verbais, mas codificam igualmente informação relativa à forma como essas situações se desenrolam no tempo. O autor demonstra, assim, que diferentes verbos, ou, mais precisamente, diferentes predicações verbais, impõem esquemas temporais distintos às eventualidades que descrevem, revelando propriedades como duração, delimitação e dinâmica interna. É neste enquadramento que Vendler introduz uma tipologia aspetual das eventualidades, amplamente usada em trabalhos posteriores.

Embora o autor apresente a tipologia de verbos, na verdade, a proposta de Vendler incide sobre predicações e não exclusivamente sobre verbos isolados. A classificação aspetual proposta por Vendler depende da interação entre o verbo e os seus argumentos, em particular o objeto direto, bem como outros elementos da predicação. Nesta perspetiva, o tempo encontra-se, em parte, lexicalmente incorporado nos verbos, mas a determinação do valor aspetual resulta da composição sintático-semântica da frase.

A tipologia de Vender assenta em critérios fundamentais, nomeadamente: (i) a compatibilidade com a forma progressiva, (ii) a existência ou não de um ponto terminal

inerente à eventualidade e (iii) a duração temporal da situação descrita. Com base nesses critérios, Vendler distingue quatro classes aspetuais fundamentais: *Activities*, *Accomplishments*, *Achievements* e *States*.

As *Activities* correspondem a predicções dinâmicas e durativas, que não possuem um ponto final inerente. Trata-se de situações homogêneas e atéticas, suscetíveis de se prolongarem indefinidamente, como em *O João correu em torno da praça*. Estas predicções são compatíveis com o progressivo e respondem a perguntas do tipo *Durante quanto tempo...?*

Os *Accomplishments* descrevem eventualidades igualmente dinâmicas e durativas, mas que pressupõem uma culminação ou fronteira final, sem a qual a situação não pode ser considerada concluída. São, por isso, predicções télicas e não homogêneas exemplos como *O João pintou um quadro*. Estas construções respondem tipicamente a perguntas do tipo *Quanto tempo demorou a...?*, refletindo a orientação para um resultado final.

Os *Achievements* designam eventos pontuais, caracterizados por uma mudança de estado que ocorre num instante específico, sem extensão temporal interna. São predicções dinâmicas, não durativas e télicas, exemplo como *O João venceu a corrida*. Embora possam ser localizadas num intervalo temporal, a eventualidade em si corresponde a um ponto no tempo.

Por fim, os *States* exprimem situações não dinâmicas, homogêneas e atéticas, que se mantêm verdadeiras ao longo de um intervalo temporal sem implicarem qualquer desenvolvimento interno ou culminação. Exemplos típicos incluem predicções como *O António está feliz*. Os estados são incompatíveis, em regra, com a forma progressiva e não descrevem mudanças ou eventos.

Esta relação diferenciada com o tempo é explicitada por Vendler através das seguintes formulações, que evidenciam o modo como cada classe se distribui temporalmente:

For activities: A was running at time t means that time instant t is on a time stretch throughout which A was running.

For accomplishments: A was drawing a circle at t means that t is on the time stretch in which A drew that circle.

For achievements: A won a race between t1 and t2 means that the time instant at which A won that race is between t1 and t2.

For states: A loved somebody from t1 to t2 means that at any instant between t1 and t2 A loved that person.

(Vendler, 1967, pp. 25-26)

Um elemento central da proposta de Vendler é o reconhecimento de que a classificação aspetual não é fixa ao nível lexical. Um mesmo verbo pode integrar classes aspetuais distintas, consoante o contexto sintático e semântico em que ocorre. O verbo *think*, por exemplo, pode funcionar como um predicado dinâmico, compatível com o progressivo, quando descreve um processo mental em curso, como em *He is thinking about Jones*, ou como um predicado estativo, quando exprime uma crença ou atitude proposicional estável, como em *He thinks that Jones is a rascal*.

A tipologia de Vendler constitui um contributo decisivo para a compreensão da dimensão aspetual das predicções, ao mostrar que a interpretação temporal dos enunciados resulta da interação entre propriedades lexicais, estrutura argumental e mecanismos gramaticais, influenciando de forma duradoura o desenvolvimento da semântica do aspeto.

4.3.2. Moens & Steedman (1988)

A proposta ontológica de Moens & Steedman (1988) para a classificação das eventualidades distingue-se por não assentar exclusivamente em primitivos temporais, como instantes ou intervalos, integrando igualmente noções de natureza concetual e discursiva, em particular a noção de contingência. Esta noção visa captar a forma como os eventos se organizam e se relacionam entre si no discurso, indo além de uma simples ordenação cronológica e incorporando fatores como expectativas, planeamento e relações experiencialmente motivadas entre situações. Moens & Steedman (1988, p. 20) ilustram este conceito com o seguinte exemplo:

(112) When John left , Sue cried.

Neste exemplo, a relação entre os dois eventos não implica necessariamente uma relação causal direta baseada no conhecimento do mundo, nem uma ordenação

temporal estrita. Como observa Leal (2009), trata-se de uma relação de contingência, entendida como uma ligação fraca e não transitiva entre eventos, distinta da causalidade forte. A contingência refere-se, assim, ao modo como os eventos se encontram interligados no discurso com base na experiência ou nas expectativas do falante, sem que exista obrigatoriamente uma relação causal lógica entre eles. Esta conceção contrasta com abordagens como a do *Penn Discourse Treebank* (PDTB), nas quais relações deste tipo tendem a ser interpretadas como relações causais.

A ausência de uma relação clara de dependência causal ou temporal linear pode observar-se no exemplo seguinte de Moens & Steedman (1988, p. 21), frequentemente considerado marginal ou infeliz do ponto de vista interpretativo:

(113) When my car broke down, the sun set.

Neste caso, os dois eventos são independentes e não estabelecem uma relação de contingência plausível, o que compromete a aceitabilidade discursiva do enunciado. Tal facto mostra que *when* nem sempre exprime um valor temporal puro ou causal, podendo, em certos contextos, limitar-se a fornecer um ponto de referência discursivo, nos termos propostos por Reichenbach (1947), sem garantir por si só a coerência relacional entre os eventos.

Com base nestas observações, Moens & Steedman (1988) defendem que a estruturação das eventualidades no discurso é influenciada por fatores como planeamento, previsibilidade, intencionalidade e expectativas sociais. A noção de contingência revela-se, assim, particularmente adequada para captar a diversidade de relações possíveis entre eventos, permitindo explicar por que razão certas sequências são interpretativamente naturais, enquanto outras resultam infelizes.

É neste contexto que os autores introduzem a noção de Núcleo Aspectual (*Aspectual Nucleus*), entendido como uma estrutura elementar que organiza as eventualidades em torno de três fases interligadas: processo preparatório, culminação e estado consequente. O processo preparatório corresponde a uma fase durativa, sem ponto terminal inerente, a culminação consiste numa transição pontual marcada por

uma mudança de estado relevante e o estado consequente corresponde ao estado resultante dessa mudança, estando contingentemente associado à culminação (Moens & Steedman, 1988).

Partindo da tipologia de Vendler (1967), que distingue *States*, *Activities*, *Accomplishments* e *Achievements*, Moens (1987) propõe uma reorganização das eventualidades com base nos critérios de pontualidade, extensão temporal e existência de estado consequente. Deste modo, distingue duas grandes categorias: eventos e estados. Os eventos incluem culminações, pontos, processos e processos culminados, caracterizando-se por apresentarem início e/ou fim definidos, enquanto os estados correspondem a situações com extensão temporal indefinida e sem dinâmica interna.

Quanto à articulação entre as classes aspetuais e o Núcleo Aspetual, a culminação envolve a fase de culminação e o estado consequente. O ponto corresponde exclusivamente à fase da culminação. O processo centra-se na fase preparatória. O processo culminado integra as três fases do Núcleo Aspetual. A Tabela seguinte sistematiza as características destas classes.

Tabela 11 - Categorias dos eventos e estados

	EVENTS		STATES
	atomic	extended	
+conseq	CULMINATION recognize, spot, win the race	CULMINATED PROCESS build a house, eat a sandwich	understand, love, know, resemble
-conseq	POINT hiccup, tap, wink	PROCESS run, swim, walk, play the piano	

Fonte: (Moens & Steedman, 1988, p. 17)

Apresentamos de seguida as definições destas classes ilustradas por exemplos de Moens & Steedman (1988). Assim, a culminação corresponde a eventos pontuais que implicam uma mudança de estado significativa, originando um estado consequente contingentemente relacionado com o evento principal (*Harry reached the top*). O ponto

designa eventos indivisíveis, sem duração interna nem estado consequente relevante (*John hiccupped*). O processo caracteriza eventos durativos, mas sem culminação inerente (*Harry climbed*). O processo culminado descreve eventos durativos que se orientam para uma culminação específica, associada a uma mudança de estado (*Harry climbed to the top*). Os estados descrevem situações estáveis, não dinâmicas, podendo subdividir-se em estados lexicais, habituais, consequentes e progressivos, como exemplificado a seguir.

(114) John loves Mary. (estado lexical)

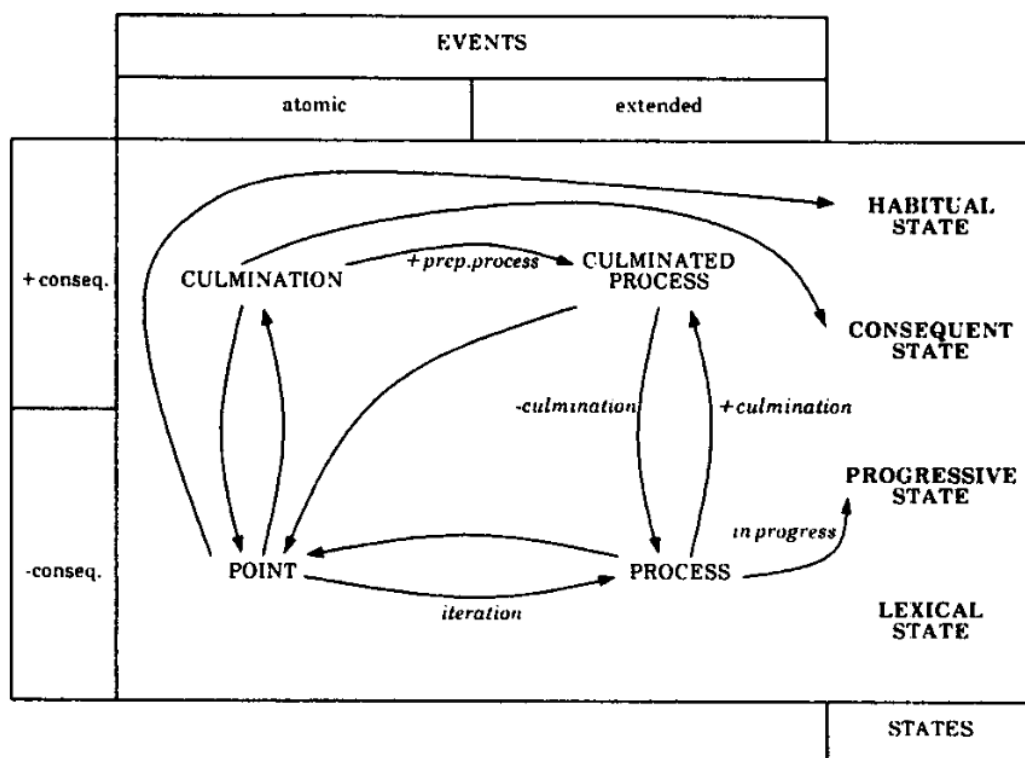
(115) Alice plays tiddlywinks for Scotland. (estado habitual) (Moens, 1987, p. 60)

(116) Harry is at the top. (estado consequente) (Moens & Steedman, p. 17)

(117) The president is speaking. (estado progressivo) (Moens & Steedman, p. 13)

Apesar de distintas, estas classes não constituem categorias rígidas. Um contributo central da proposta de Moens & Steedman (1988) é a explicitação dos mecanismos de coerção aspetual, que permitem transições entre diferentes tipos de eventualidades mediante a intervenção de operadores aspetuais, tempos verbais, advérbios temporais ou auxiliares. A representação das transições aspetuais possíveis na rede proposta por Moens e Steedman está sistematizada na Figura seguinte.

Figura 10 - Rede de transição aspetual



Fonte: (Moens & Steedman, 1988, p. 18)

No que diz respeito aos operadores aspetuais que podem desencadear mudanças nesta Rede Aspetual, o progressivo, por exemplo, denota uma perspetiva durativa e não pontual. Quando aplicado a eventos pontuais, desencadeia coerção aspetual, reinterpretando-os como processos (Moens & Steedman, 1988, p. 17):

(118) Harry hiccupped.

(119) Harry was hiccupping.

Quando aplicado a processos culminados, o progressivo elimina a culminação e focaliza exclusivamente o processo preparatório (Moens & Steedman, 1988, p. 17):

(120) Roger ran a mile.

(121) Roger was running a mile.

O perfectivo, por contraste, exige a presença de uma culminação e centra-se na fase culminante da eventualidade.

Por fim, Moens & Steedman (1988, p. 20) recorrem a advérbios temporais do tipo *durante x tempo* e *em x tempo* como testes aspetuais. Os primeiros são compatíveis com processos e podem transformar eventos télicos em atélicos. Os segundos ocorrem tipicamente com processos culminados e podem transformar eventos atélicos em eventos télicos.

Estes mecanismos confirmam que a interpretação aspetual resulta da interação dinâmica entre propriedades lexicais, operadores gramaticais e estrutura discursiva, sendo a noção de contingência central para compreender a organização temporal dos eventos no discurso.

4.3.3. Cunha (2004; 2013)

Os trabalhos de Cunha (2004; 2013) permitem verificar que as propostas feitas por Vendler (1967) e Moens & Steedman (1988) são adequadas à representação do valor aspetual das predicções ao português europeu.

Em Cunha (2013), o autor analisa as classes aspetuais em português europeu com base nas propriedades de dinamicidade, duratividade, telicidade e homogeneidade, seguindo autores como Vendler (1967) e Moens & Steedman (1988). Retoma a noção de situação dinâmica como se referindo a eventos que se desenvolvem através de fases sucessivas, implicando uma alteração do estado inicial ao longo da linha temporal e de situação não dinâmica ou estática como não envolvendo esse tipo de mudança. Quanto à duratividade e telicidade, apresenta a primeira como dizendo respeito à extensão do evento ao longo de um intervalo de tempo, enquanto a telicidade como caracterizando eventos que apresentam um limite terminal definido, conceitos já discutidos por Vendler (1967) e Moens & Steedman (1988). A homogeneidade, por sua vez, é definida como se referindo a situações em que cada subevento mantém as mesmas propriedades do evento principal.

Neste enquadramento teórico, Cunha (2013) adapta a proposta de Moens & Steedman (1988) aos exemplos e ao português, como se pode observar na Figura seguinte.

Tabela 12 - Classes aspetuais básicas

	Dinâmico		Não dinâmico
	EVENTOS		ESTADOS
	Pontual	Durativo	amar, compreender, estar sentado, saber
Télico	CULMINAÇÕES <i>avistar, ganhar a corrida, reconhecer</i>	PROCESSOS CULMINADOS <i>comer uma sandes, construir a casa</i>	
Não télico	PONTOS <i>espirrar, pestanejar, tossir</i>	PROCESSOS <i>andar, correr, nadar, tocar piano</i>	

Fonte: (Cunha, 2013, P. 591)

Com base nas propriedades apresentadas anteriormente, Cunha (2013) classifica as situações ou eventualidades em cinco categorias: estado, processo, processo culminado, culminação e ponto. Atendendo à complexidade inerente à distinção entre estados e eventos, o autor apresenta um conjunto de critérios ou testes linguísticos que permitem clarificar essa diferença na linha dos autores já referidos, como Dowty (1979). Ao contrário dos eventos, os estados apresentam uma leitura preferencialmente puramente temporal, sendo frequentemente expressos através do Presente do Indicativo. Enquanto os estados transmitem informação sobre o momento presente, os eventos tendem a apresentar leituras habituais. Além disso, os eventos durativos são compatíveis com operadores aspetuais como *parar de* ou *acabar de*, enquanto os estados não o são. Por outro lado, as situações estativas podem ocorrer com o operador aspetual *passar a*, enquanto os eventos se combinam preferencialmente com o operador *começar a*.

Com o objetivo de assegurar uma identificação mais precisa da distinção entre eventos e estados, o autor aplica ainda outros testes linguísticos já propostos por autores como Dowty (1979) aos dados em português, nomeadamente: (i) a

possibilidade de ocorrência em orações imperativas; (ii) a possibilidade de integração em orações subordinadas a predicadores de comando ou de persuasão; (iii) a possibilidade de coocorrência com advérbios volitivos; e (iv) a possibilidade de o sintagma verbal constituir o foco de uma oração pseudoclivada.

Para testar e distinguir diferentes tipos de eventos, o autor propõe diagramas baseados nas estruturas *durante x tempo* (processo), *em x tempo* (processo culminado) e *às x tempo* (culminação). As duas primeiras representações retomam a proposta de Moens & Steedman (1988), sendo utilizadas para testar a transição entre processos e processos culminados.

Um elemento particularmente relevante da proposta de Cunha (2013) é a distinção entre estados estáveis vs. estados episódicos e estados faseáveis (individual-level vs. stage-level, Carlson (1977)) vs. estados não faseáveis. Os estados estáveis exprimem propriedades intrínsecas e duradouras dos indivíduos, como “ser feliz, ser inteligente, ter olhos azuis, viver em Lisboa, gostar de linguística, ser alto, saber inglês” (Cunha, 2013, p. 595). Os estados episódicos, por sua vez, referem-se a propriedades transitórias, como “estar contente, estar bêbado, estar vestido, ter frio, ter fome, estar doente, ter dores de cabeça” (Cunha, 2013, p. 595). Por fim, a distinção entre estados faseáveis e não faseáveis, inicialmente proposta em Cunha (1998) e retomada em Cunha (2004), constitui uma proposta relevante para caracterizar estados com comportamentos de eventos. De facto, os estados faseáveis aproximam-se do comportamento dos eventos, por serem mais suscetíveis de variação em determinados contextos, como “ser sincero, viver em Paris, ser agressivo, gostar de linguística, ser preguiçoso” (Cunha, 2013, p. 595).

Cunha (2013) revê ainda fatores linguísticos que podem influenciar a determinação das classes aspetuais. O verbo, por si só, pode contribuir para as propriedades aspetuais da frase, mas outros constituintes também desempenham um papel relevante. O autor mostra que também em português o sujeito e o complemento direto podem influenciar a interpretação télica ou atélica, ou durativa e não durativa, da eventualidade, enquanto os sintagmas preposicionais espaciais podem igualmente influenciar a leitura aspetual das predicções.

4.3.4. Conclusão

Neste capítulo, abordaram-se os conceitos de Tempo e de Aspeto, procedendo-se a uma revisão de algumas das principais propostas teóricas para o tratamento destas duas categorias que, como se verificou, são interdependentes. Foram, assim, discutidos elementos fundamentais para a descrição das propriedades temporais e aspetuais das sequências de tempos que ocorrem no discurso.

Uma vez que as relações causais se estabelecem entre duas situações e que, tal como foi mostrado no Capítulo 2, se encontram frequentemente associadas a uma relação de sucessividade temporal, a caracterização das relações temporais, bem como das propriedades aspetuais das situações envolvidas, constitui uma dimensão central para a análise da causalidade. Deste modo, tanto a caracterização temporal como a caracterização aspetual das situações serão dois dos parâmetros fundamentais a considerar na análise dos dados que apresentaremos no capítulo 6.

5. Esquemas de anotação semântica de informação temporal e causal

A temporalidade e a causalidade constituem duas temáticas amplamente estudadas ao longo dos anos, em particular no que diz respeito às suas construções, inferências e aos elementos linguísticos envolvidos no processamento da linguagem natural, sobretudo nos domínios da compreensão de eventos e da análise do discurso escrito. De acordo com Amorim et al.,(2025), a representação e o processamento da dimensão temporal, bem como da causalidade, têm sido objeto de investigação em diversos campos nos últimos anos, tanto em Linguística como em Processamento de Linguagem Natural. Apesar disso, os trabalhos existentes tendem a centrar-se predominantemente numa destas dimensões em detrimento da outra, sendo ainda reduzido o número de estudos que abordam simultaneamente a temporalidade e a causalidade.

Tal como é sublinhado por Fernandes et al. (2025a), a anotação manual assume um papel central quer na investigação em Linguística quer no desenvolvimento de aplicações em Processamento de Linguagem Natural. Para além de permitir a análise sistemática de fenómenos linguísticos, a anotação fornece dados de referência fiáveis, indispensáveis à consolidação e operacionalização de teorias linguísticas, ao tornar explícitas categorias, distinções e generalizações que podem ser testadas empiricamente. Neste contexto, a definição de esquemas de anotação bem fundamentados revela-se crucial. Fernandes et al. (2025b) definem esquemas de anotação como conjuntos estruturados de etiquetas de natureza descritiva e analítica que são atribuídos aos dados linguísticos ao longo do processo de anotação. Esta atribuição é orientada por diretrizes previamente estabelecidas, nas quais se especificam as etiquetas a utilizar, as suas propriedades/ atributos, as unidades de anotação relevantes (como palavras, expressões, orações) e os procedimentos a seguir. Para garantir a coerência do processo, é essencial que tanto as etiquetas como as unidades sejam definidos de forma operacional e inequívoca, de modo a promover o consenso entre anotadores humanos. Ao articular princípios teóricos com exigências práticas, estes esquemas permitem representar fenómenos linguísticos complexos de

forma rigorosa, sem perder de vista as limitações e necessidades dos anotadores humanos. Assim, a anotação manual, orientada por esquemas sólidos, emerge como um elemento-chave no estudo em Linguística, e em particular em Semântica.

No âmbito de esquemas de anotação semântica, há algumas propostas que integram etiquetas e atributos que podem representar as propriedades temporais, aspetuais e discursivas. Entre essas propostas, destacam-se os trabalhos de Pustejovsky et al. (2003b), Costa & Branco (2012) e Silvano et al. (2024) no domínio da temporalidade; de Girju et al. (2007), Dunietz et al. (2015; 2017) e Tan et al. (2022) no domínio da causalidade; e de Bethard et al. (2008) e Mirza & Tonelli (2014), que integram conjuntamente a temporalidade e a causalidade.

Neste capítulo, apresenta-se, uma síntese de alguns desses trabalhos, descrevendo os esquemas de anotação propostos, os respetivos corpora anotados e as contribuições mais relevantes para este domínio de investigação, no âmbito da Semântica.

5.1. Esquemas de anotação causal

A causalidade constitui um dos princípios fundamentais do pensamento humano, permitindo aos indivíduos formular, compreender e interpretar fenómenos naturais e sociais no seu quotidiano. Como já referimos nos capítulos anteriores, ao longo dos anos, numerosos estudos têm procurado compreender este princípio, salientando o seu papel central na forma como os seres humanos constroem julgamentos sobre a realidade que os rodeia. A investigação em torno da causalidade continua, assim, a suscitar interesse em diferentes áreas do conhecimento, em particular no domínio da Semântica, trazendo contributos relevantes para a área da linguística, enquanto disciplina dedicada à descrição do funcionamento da língua.

Um dos trabalhos de referência na análise das relações causais, já discutido no subcapítulo 2.5., é o estudo de Prasad et al. (2008), no qual é apresentado um corpus pioneiro dedicado à anotação de relações discursivas em textos do *Wall Street Journal*. Este corpus inclui mais de 6.000 ocorrências anotadas de relações de causalidade, enquadradas na hierarquia semântica da classe CONTINGENCY (cf. Figura 9, subcapítulo 2.5.).

Retomamos aqui a distinção feita pelos autores entre relações discursivas explícitas e implícitas. As relações causais explícitas correspondem a casos em que a relação é assinalada por conectores discursivos, como *because* ou *as a result*. Estes conectores são classificados em subordinativos, coordenativos, adverbiais causais e estruturas causais modificadas. Nestas situações, a anotação segue o vínculo sintático do conector: o argumento ao qual o conector causal se encontra gramaticalmente associado é identificado como Arg2, geralmente interpretado como a causa, enquanto o outro argumento é anotado como Arg1, correspondendo, na maioria dos casos, à consequência. A ordem linear dos argumentos pode variar, dando origem a três configurações principais: Arg1–Arg2, Arg2–Arg1 ou Arg2 encaixado em Arg1 (Prasad et al., 2008), correspondendo, respetivamente, às estruturas Arg1–Conector+Arg2; Conector+Arg2–Arg1; e Arg1 com Conector+Arg2 embutido. As relações causais implícitas, por sua vez, são inferidas entre orações ou frases adjacentes na ausência de um conector causal explícito. Nestes casos, a anotação baseia-se exclusivamente na ordem linear do texto, sendo o primeiro segmento identificado como Arg1 e o segmento seguinte como Arg2.

Outro contributo relevante para o estudo da causalidade é o trabalho de Dunietz et al. (2015), intitulado *Annotating Causal Language Using Corpus Lexicography of Constructions*. Os autores partem da constatação de que “there is no one standard representation scheme that can handle all of these types of semantics, making it difficult to analyze and extract causal relationships in a coherent, comprehensive manner” (p. 188). Esta dificuldade decorre do facto de a causalidade ser entendida como um construto psicológico utilizado para explicar o mundo, não correspondendo de forma direta nem à realidade empírica nem às estruturas linguísticas que a descrevem: “causation is a psychological construct that we use to explain the world around us, and it does not perfectly match either empirical reality or the language we use to describe it” (Dunietz et al., 2015, p. 188). Os autores referem ainda que a causalidade encontra-se, assim, intimamente relacionada com outras dimensões semânticas, como a temporalidade, os contrafactuais, a factualidade e a negação.

Para lidar com esta complexidade, Dunietz et al. (2015) propõem um novo esquema de anotação para a causalidade, fundamentado nos princípios da Gramática de Construções (CxG) (Fillmore et al., 1988; Goldberg, 1995). A CxG concebe as construções como unidades fundamentais da linguagem, definidas como formas linguísticas de complexidade variável, desde morfemas até padrões léxico-sintáticos estruturados. Neste enquadramento, a linguagem causal é entendida como “any construction which presents one event, state, action, or entity as promoting or hindering another, and which includes at least one lexical trigger” (Dunietz et al., 2017, p. 97).

A proposta de Dunietz et al. (2015) privilegia, assim, a análise das expressões linguísticas de causalidade, abstraindo-se das questões filosóficas relativas à causalidade no mundo real. O estudo centra-se em contextos em que as relações causais são expressas explicitamente em inglês, partindo do pressuposto de que a causalidade é um dos tipos mais relevantes de relação semântica a ser extraído. Segundo os autores, a causalidade está omnipresente na linguagem, sendo frequentemente expressa por verbos. Contudo, defendem que “causal relationships are not captured well by standard semantic representations. The linguistic expression of causal relations varies greatly (...), ranging from verbal propositions to discourse relations to arbitrarily complex constructions” (Dunietz et al., 2015, p. 188).

Neste contexto, Dunietz et al. (2015) optam por anotar exclusivamente relações causais explícitas, excluindo casos em que a causalidade é expressa de forma implícita, isto é, sem a presença de um marcador lexical. São igualmente excluídos verbos causativos que apenas indicam o resultado da causalidade, bem como conectores que introduzem relações causais vagas ou inespecíficas. Por fim, são excluídos da anotação casos ambíguos, como orações introduzidas por *after*, cuja interpretação causal depende do conhecimento extralinguístico e não de um marcador causal explícito. Estas restrições refletem a opção dos autores por uma abordagem estritamente linguística à anotação da causalidade.

Com base nesta delimitação, o estudo centra-se na identificação de três componentes fundamentais da relação causal: o conector, entendido como a marca lexical que sinaliza a relação, independentemente do padrão linguístico; a causa,

correspondente a eventos, ações ou situações expressas sob a forma de orações ou frases; e o efeito, que designa um resultado ou estado de coisas igualmente expresso por uma oração completa. A partir destes critérios, são considerados quatro tipos de causalidade: *consequence*, *motivation*, *purpose* e *inference*, sendo anotados os segmentos textuais correspondentes ao conetor causal, ao argumento de causa e ao argumento de efeito no corpus BECauSE 1.0 (Dunietz et al., 2015).

O trabalho de Dunietz et al. (2017) apresenta uma versão alargada deste corpus, designada BECauSE 2.0, com anotações exaustivas das relações causais. Os autores concetualizam a causalidade como redes causais, nas quais fenómenos causam, possibilitam ou impedem outros fenómenos, sublinhando que “causal networks – phenomena causing, enabling, or preventing others. Accordingly, the language we use is full of references to cause and effect” (Dunietz et al., 2017, p. 95). As proposições causais são consideradas particularmente relevantes para aplicações orientadas para a Semântica, sobretudo em domínios como a finança e a biologia, onde a sua interpretação pode apoiar processos de tomada de decisão.

O BECauSE 2.0 pode ser visto como uma adaptação da abordagem de anotação múltipla do PDTB, embora não se restrinja exclusivamente às relações discursivas. Pelo contrário, visa capturar todos os significados expressos pelas construções causais, anotando cada instância com a totalidade dos sentidos que transmite. Este corpus difere da primeira versão em termos de escopo e de riqueza da anotação linguística usada para lidar com linguagem causal. Enquanto o BECauSE 1.0 se centra em definir e anotar instâncias de linguagem causal explícita, capturando construções que expressam causalidade no texto e oferecendo um esquema uniforme para representar causas e efeitos em termos linguísticos num corpus, o BECauSE 2.0 estende e melhora essa versão original ao incluir não só as expressões causais exaustivamente anotadas, mas também sete tipos de relações semânticas que frequentemente coocorrem com ou se sobrepõem à causalidade (como temporalidade, correlação, hipotéticos, obrigação/permissão, criação/terminação, extremidade/suficiência e contexto). Isso permite uma anotação mais detalhada e precisa das construções linguísticas, refletindo

melhor a complexa interação entre causalidade e outras relações semânticas no uso real da língua.

No esquema de BeCause 2.0, distinguem-se três tipos principais de relações causais: Consequência, Motivação e Propósito, sendo eliminada a categoria de Inferência previamente considerada. O esquema passa ainda a diferenciar entre causação positiva (FACILITAR) e causação inibitória (INIBIR).

Enquanto extensão do BECauSE 1.0, o BECauSE 2.0 é composto por 59 artigos, 47 documentos selecionados aleatoriamente, 679 frases provenientes de transcrições de audiências, 10 notícias com um total de 547 frases e 2 jornais dos quais foram extraídas 82 frases. O corpus totaliza 5.380 frases, incluindo 1.803 instâncias anotadas de linguagem causal, das quais 1.634 (90,7%) apresentam explicitamente argumentos de causa e efeito. Cerca de um terço destas instâncias envolve relações semânticas sobrepostas, sendo ainda registadas 583 anotações de relações sobrepostas não causais.

Outro contributo relevante é o estudo de Tan et al. (2022), no qual se defende que “causality is a core cognitive concept and appears in many natural language processing (NLP) works that aim to tackle inference and understanding” (p. 2298). Neste trabalho, a causalidade é definida como uma relação semântica entre dois argumentos, causa e efeito, em que a ocorrência de um provoca a ocorrência do outro. A relação causal envolve, assim, um facto que funciona como argumento de causa e a sua consequência, designada como argumento de efeito.

Tan et al. (2022) identificam uma discrepância entre as diretrizes de anotação existentes para a causalidade de eventos e os corpora convencionais, propondo, para a resolver, diretrizes alinhadas com esquemas linguísticos e de causalidade de eventos previamente estabelecidos. Os autores observam que muitas abordagens se limitam a relações explícitas ou a argumentos expressos por orações completas. Em resposta, propõem um esquema de anotação baseado em eventos, classificando as frases como causais ou não causais. Nas frases causais, a causalidade pode ser explícita ou implícita, sendo obrigatória a presença simultânea de causa e efeito; nas frases não causais, a ausência de um ou de ambos os argumentos impede a atribuição de uma relação causal.

Para tornar a anotação mais visível, o esquema recorre a diferentes cores para representar os elementos envolvidos: causas em rosa, efeitos em verde e sinais causais em amarelo. A identificação dos eventos baseia-se na definição de evento do TimeML, sendo a noção de causalidade inspirada na definição de Contingency do PDTB-3, segundo a qual “one argument provides the reason, explanation or justification for the situation described by the other” (PDTB-3, p. 19). Os autores distinguem quatro sentidos causais: cause, purpose, condition e negative-condition. Para validar a existência de uma relação causal, recorrem a cinco testes propostos por Grivaz (2010) e Dunietz et al. (2017): *Why*, *Temporal order*, *Counterfactual*, *Ontological asymmetry* e *linguistic*. O teste *Why* (Porquê) verifica se faz sentido perguntar por que razão o efeito ocorreu, sendo a resposta a causa. O teste da *Temporal order* (Ordem temporal) confirma que a causa precede o efeito. O teste *Counterfactual* (Contrafactual) avalia se a ausência da causa implicaria a não ocorrência do efeito. O teste de *Ontological asymmetry* (Assimetria ontológica) estabelece o caráter não simétrico da relação causal. Por fim, o teste *Linguistic* (Linguístico) verifica se a relação pode ser reformulada como “X causa Y”.

Com base nestes critérios, foi criado o Causal News Corpus (CNC), composto por 869 notícias provenientes de diferentes fontes e períodos. No total, foram identificadas 3.559 frases com eventos em notícias em língua inglesa relacionadas com protestos, rotuladas “with binary labels indicating whether each sentence contains a causal relationship or not” (Tan et al., 2022, p. 2299). A anotação foi realizada por dois anotadores, sendo os resultados avaliados aleatoriamente por um supervisor e posteriormente corrigidos de forma semiautomática.

5.2. Esquemas de anotação temporal

No conjunto das propostas de anotação temporal, destaca-se o trabalho de Pustejovsky et al. (2003b), no qual é desenvolvido um dos esquemas mais amplamente utilizados para a anotação e a análise da informação temporal em textos. Os autores propõem o esquema de anotação TimeML, que resulta da evolução do sistema TIMEX2 (Ferro et al., 2001; Setzer, 2001) e visa representar, de forma sistemática, eventos, expressões temporais e as relações estabelecidas entre eles. Esta proposta está na base da ISO

24617-1 (2012), *Language resource management-Semantic annotation framework: Part 1- Time and events*.

De acordo com Pustejovsky et al. (2003b), a abrangência e a flexibilidade do TimeML permitem a anotação de quatro tipos principais de entidades e relações: EVENT, TIMEX3, SIGNAL e LINK.

A etiqueta EVENT destina-se à anotação de eventualidades, incluindo eventos pontuais ou durativos, bem como certos estados ou circunstâncias. Contudo, nem todos os predicados estativos são anotados como EVENT: apenas os são aqueles que contrastam com outros estados ou que introduzem uma mudança relevante para a interpretação temporal do discurso. Estados considerados comuns e sem impacto na interpretação temporal não são anotados no esquema. A etiqueta EVENT inclui atributos linguísticos como a classe aspetual, o tempo verbal, o aspeto e a classe gramatical, entre outros.

A etiqueta TIMEX3 é utilizada para anotar expressões temporais, classificando-as em três categorias principais: expressões temporais totalmente especificadas, expressões temporais subespecificadas e durações (Pustejovsky et al., 2003b). Esta anotação permite representar tanto referências temporais absolutas como relativas ao contexto discursivo.

A etiqueta SIGNAL corresponde a palavras ou expressões que explicitam relações temporais, aspetuais ou de subordinação. Incluem-se neste grupo preposições temporais, conetores temporais e elementos subordinantes. Para além disso, esta etiqueta pode também ser utilizada para anotar quantificadores temporais e certos fenómenos de polaridade frásica (Pustejovsky et al., 2003b).

A etiqueta LINK codifica as relações estabelecidas entre eventos e entre eventos e expressões temporais. Estas relações incluem, por um lado, relações temporais, como identidade, anterioridade e posterioridade (Pustejovsky et al., 2003b; Pustejovsky et al., 2010). Incluem ainda relações de subordinação, como INTENSIONAL, FACTIVE e CONDITION e, ainda, relações aspetuais, que dizem respeito às ligações entre eventos aspetuais e os seus argumentos (Pustejovsky et al., 2003b).

Para além destas quatro etiquetas centrais, o esquema TimeML disponibiliza mecanismos que permitem representar expressões temporais dependentes do contexto sem necessidade de atribuição explícita de um valor absoluto, conforme previsto posteriormente na norma ISO 24617-1 (2012). O esquema reconhece diferentes tipos de eventos, incluindo verbos flexionados e não flexionados, adjetivos com valor estativo ou eventivo, nominalizações, orações predicativas e sintagmas preposicionais com valor de evento ou estado. Permite ainda estabelecer diferentes tipos de ligações entre eventos e expressões temporais, nomeadamente relações de ancoragem temporal, de ordenação temporal e de subordinação narrativa ou discursiva (Pustejovsky et al., 2003b).

Devido à sua relevância, o TimeML tem sido amplamente utilizado em estudos de análise discursiva e temporal em diversas línguas, como o inglês, o francês e o português (Pustejovsky et al., 2003b; Costa & Branco, 2012; Silvano et al., 2024). Atendendo às especificidades estruturais de cada língua, foram desenvolvidas adaptações do esquema original, das quais resultaram versões como o KTimeML (Im et al., 2009) e o It-TimeML (Caselli et al., 2011).

O estudo apresentado em Pustejovsky et al. (2003a), intitulado *The TimeBank Corpus*, tem como base um conjunto de 300 artigos provenientes de diferentes agências noticiosas. Nesse corpus, foram anotados 7.571 eventos, destacando-se os atributos *class*, *tense* e *aspect*, sendo a classe de ocorrência a mais frequente. Foram igualmente identificadas 1.423 expressões temporais, caracterizadas, entre outras propriedades, pelos atributos *type* e *FunctionInDocument*, predominando as expressões do tipo *data*. Adicionalmente, registaram-se 2.212 sinais temporais e um total de 38.242 ligações, das quais 62% corresponderam a relações temporais.

No contexto do português europeu, os trabalhos de Costa & Branco (2010; 2012) adaptaram o esquema TimeML para a anotação de informação temporal em textos traduzidos do corpus TimeBank, dando origem ao TimeBankPT, o primeiro corpus com foco específico na análise temporal nesta variedade do português. As anotações mantiveram-se, em larga medida, compatíveis com o esquema original, embora tenham

sido introduzidas adaptações motivadas por diferenças estruturais e linguísticas entre o inglês e o português europeu.

Importa ainda referir a ISO-TimeML, integrada na norma ISO 24617, dedicada à gestão de recursos linguísticos. Esta norma define uma arquitetura geral para a anotação semântica, organizada em diferentes partes correspondentes a distintos níveis de análise, como o temporal (ISO 24617-1), o referencial (ISO 24617-4), o espacial (ISO 24617-7), as funções semânticas (ISO 24617-9) e as relações discursivas (ISO 24617-8), entre outros. No que diz respeito especificamente à anotação temporal, destaca-se a parte ISO 24617-1, que define:

a formalized XML-based markup language called ISO-TimeML, with a systematic way to extract and represent temporal information, as well as to facilitate the exchange of temporal information, both between operational language processing systems and between different temporal representation schemes (ISO 24617-1, p. 1).

Embora inspirada no TimeML (Pustejovsky et al., 2003b), a ISO-TimeML introduz um conjunto de alterações e ajustes. Entre estes, incluem-se a uniformização dos atributos de identificação das etiquetas (por exemplo, *eid*, *tid* e *sid* substituídos por *xml:id*), a eliminação de certos elementos e a introdução de novos atributos com *pred* (predicado), entre outros aspetos. Estruturalmente, a ISO-TimeML organiza-se em duas componentes principais: a Estrutura de Entidades e a Estrutura de Ligações. A anotação temporal envolve, assim, entidades como *Timex3* e *Event*, bem como relações codificadas através de *TLINK* (*Temporal links*), abrangendo propriedades de eventos, relações temporais entre eventos e expressões temporais, relações de subordinação e relações aspetuais (ISO 24617-1, 2012).

Partindo da ISO 24617 – *Language Resource Management: Semantic Annotation Framework* (SemAF), Silvano et al. (2021) e Leal et al. (2022) propõem um esquema de anotação multicamada, o *Text2Story annotation scheme*. Este esquema procura harmonizar diferentes partes desta norma, combinando, de forma integrada, os componentes temporal (ISO 24617-1), referencial (ISO 24617-9), espacial (ISO 24617-7) e de papéis semânticos (ISO 24617-4). O esquema foi inicialmente desenvolvido para representar a informação semântica de narrativas jornalísticas. Tendo em conta esse objetivo, os autores procederam à exclusão de alguns atributos e tipos de ligação

previstos na norma ISO-TimeML e reorganizaram o esquema de anotação em duas grandes dimensões. Mantiveram, no entanto, as entidades EVENT (cf. Bach, 1985), com os atributos *class*, *type*, *pos*, *tense*, *aspect*, *vform*, *polarity*, *mood* e *modality*, mas com modificações em alguns atributos para dar conta das especificidades da língua; e a entidade TIME, com os atributos essenciais *type*, *value*, *temporal function* e *anchor time*. A segunda dimensão diz respeito às ligações (LINKs), englobando *temporal links* (por exemplo, *before*, *after*, *includes*, *is_included*, *during*, *identity*), *aspectual links* e *subordination links* (Silvano et al., 2021; 2023). A coerência, assim como a interoperabilidade do esquema de anotação *Text2Story* foram já testadas e validadas em diversos estudos, nomeadamente em Nunes et al. (2024), Silvano et al. (2024) e Evelin et al. (2025), que reportam resultados positivos quanto à sua aplicação em contextos de anotação manual e análise discursiva.

Com recurso ao esquema de anotação *Text2Story*, foram anotadas manualmente 117 notícias utilizando a plataforma BRAT (Stenetorp et al., 2012). O corpus resultante integra o *Text2Story Lusa dataset*, cuja anotação “*comprises entity structures (events, times, participants, measures, and spatial relations) and link structures (e.g., temporal, aspectual, subordination, objectal, spatial, and semantic role links)*” (Nunes et al., 2024). Trata-se, assim, de um corpus com um nível elevado de detalhe de anotação.

Para além disso, este esquema de anotação tem permitido investigar diferentes características das notícias, nomeadamente a organização temporal entre o nível da história da notícia e do relato da história (cf. Silvano et al., 2024) e entre os eventos e as expressões temporais que são usadas para os localizar (cf. Amorim et al., 2025).

Em síntese, TimeML, ISO-TimeML e adaptações como o esquema *Text2Story* constituem esquemas de anotação de grande relevância para a análise de fenómenos linguísticos relacionados com o tempo e o discurso, contribuindo de forma significativa para o desenvolvimento de recursos em Linguística e em Linguística Computacional.

5.3. Esquemas de anotação causal e temporal

A articulação da causalidade com a temporalidade num mesmo esquema de anotação é muito pouco frequente. Tradicionalmente, estas duas dimensões têm sido estudadas de forma independente, conduzindo ao desenvolvimento de esquemas de

anotação distintos e à criação de recursos específicos para cada uma delas. No entanto, as relações causais estão estreitamente interligadas às relações temporais estabelecidas entre os eventos, sendo frequentemente associadas a uma relação temporal em que a causa preceda o efeito (cf. Rink et al. 2010, por exemplo).

Entre os estudos que procuram integrar estas duas dimensões, destaca-se o trabalho de Bethard & Martin (2008), que procura preencher lacunas deixadas por corpora anteriores que tratavam estas relações de forma separada. O esquema de anotação proposto pelos autores caracteriza-se por ser bidimensional, permitindo a anotação simultânea de relações causais e temporais entre pares de eventos. No que diz respeito à dimensão temporal, o esquema recorre a apenas três rótulos: *before*, *after* e *No-Rel*. Relativamente à anotação causal, a abordagem inicial baseava-se numa conceptualização da causalidade assente em condições necessárias e suficientes, o que revelou algumas limitações. Para resolver este problema, os autores propuseram uma metodologia alternativa centrada na identificação de diferentes pares de eventos ligados pela conjunção *and*, uma construção frequente que pode indicar simultaneamente relações temporais e causais.

A abordagem adotada para as relações causais não se baseou apenas na presença intuitiva de causalidade no texto, mas sim numa técnica de paráfrase linguística semelhante à proposta de Bethard & Martin (2008). Os anotadores tinham de verificar se a conjunção *and* poderia ser substituída por uma expressão que explicitasse causalidade. As opções incluíam, por exemplo: CAUSAL: quando a conjunção poderia ser substituída por expressões como *and as a result*, *and as a consequence*, ou *and enabled by that*, sinalizando uma relação de causa-efeito explícita; NO-REL: quando a conjunção não podia ser parafraseada de forma causal, como em *and independently* ou *and for similar reasons*, indicando a ausência de ligação causal clara (Bethard & Martin, 2008, p. 178).

O esquema foi aplicado ao corpus *Penn Treebank*, no qual foram identificados 5.013 pares de eventos coordenados, distribuídos por 556 documentos. Deste conjunto, foi selecionado um subconjunto de 1.000 pares de eventos para análise. Quanto aos pares de eventos ligados apenas por relações temporais, foram anotadas 313

ocorrências de *before* e 16 de *after*. Quanto às relações causais, registaram-se 217 ocorrências. A análise dos dados revela que os valores temporais em frases com *and* ocorrem, em média, cerca de nove vezes por documento, enquanto as relações causais explícitas surgem com uma frequência inferior a uma ocorrência por documento, indicando que a causalidade explícita é relativamente rara em corpora jornalísticos.

Outro contributo relevante é o estudo de Mirza & Tonelli (2014), que propõe um esquema de anotação destinado à análise de relações causais explícitas entre pares de eventos e respetivos marcadores linguísticos. Este trabalho adota uma perspetiva integrada sobre a interação entre causalidade e temporalidade. Nesta proposta, a temporalidade é analisada a partir das relações entre eventos e expressões temporais, seguindo o padrão do TimeML. A causalidade lexicalizada, por sua vez, fundamenta-se na teoria da dinâmica de forças de Talmy (1985; 1988), sendo os sinais causais classificados com base em verbos de *Cause*, *Enable* e *Prevent* (Wolff, 2007). A causalidade lexicalizada é, assim, entendida como a expressão linguística explícita de relações causais através de itens lexicais específicos, em particular verbos. Esta abordagem fundamenta-se na teoria da dinâmica de forças, proposta por Talmy (1985, 1988), segundo a qual os eventos causais podem ser concetualizados como interações entre entidades dotadas de forças, intenções ou tendências opostas. Nesta perspetiva, a causalidade não é descrita apenas como uma sequência de acontecimentos no tempo, mas como o resultado da interação entre uma entidade que exerce uma força e outra que lhe resiste ou é por ela afetada.

Com base neste enquadramento concetual, Wolff (2007) propõe uma classificação dos sinais causais lexicalizados em função do tipo de interação de forças que expressam. Assim, distinguem-se verbos que codificam relações de *Cause*, nos quais uma entidade exerce uma força que produz diretamente um determinado efeito; verbos de *Enable*, que expressam situações em que uma entidade remove um obstáculo ou cria condições para que um evento ocorra; e verbos de *Prevent*, que descrevem casos em que uma entidade impede ou bloqueia a ocorrência de um evento. Esta classificação permite descrever de forma mais precisa os diferentes modos pelos quais a causalidade

é codificada lexicalmente, captando distinções semânticas relevantes entre causar, possibilitar e impedir um determinado estado de coisas.

Ao nível da anotação, o esquema contempla eventos (com os atributos *Event_Type* e *pos*), expressões temporais, sinais causais (C-SIGNAL), incluindo preposições, conjunções, conectores adverbiais e expressões integradas na oração, relações temporais (como *before*, *after* e *includes*) e relações causais explícitas, tanto no âmbito da frase, como entre frases. Tanto as relações temporais como as causais são anotadas seguindo a orientação de um evento de origem (causa) para um evento de destino (efeito).

Com base no corpus *TimeBank*, já anotado com informação temporal, composto por 183 documentos, a anotação da causalidade foi realizada com recurso à ferramenta CAT (Bartalesi Lenzi et al., 2012). Foram identificadas 319 relações causais em 109 dos 183 documentos analisados. Adicionalmente, registaram-se 171 marcas linguísticas causais explícitas em 87 documentos, incluindo verbos de afetação (*affect*, *influence*, *determine*, *change*, entre outros), verbos causativos básicos com valores de *Cause*, *Enable*, *Prevent*, construções perifrásticas de causalidade com valores de *Cause*, *Enable*, *Prevent* e *C-Signals*. Tendo em consideração que o corpus tinha 5.118 relações temporais entre eventos anotadas, destaca-se uma clara disparidade entre a frequência da temporalidade e a da causalidade explícita.

Neste enquadramento, o presente estudo distingue-se de abordagens anteriores por não se limitar à análise isolada da causalidade, integrando igualmente a dimensão temporal na análise das relações causais. Esta perspetiva permite uma caracterização mais abrangente da organização dos eventos no discurso, tendo em conta a interação entre relações causais e as suas propriedades temporais.

5.4. Conclusão

Neste capítulo, procedeu-se a uma revisão de algumas das principais propostas de esquemas de anotação da informação semântica temporal e causal. Como se pôde constatar, existe uma ampla diversidade de esquemas dedicados à anotação temporal, enquanto que os esquemas especificamente orientados para a anotação da causalidade

são menos numerosos. Ainda mais escassas são as propostas que integram, de forma sistemática, a representação conjunta das relações temporais e causais.

Embora algumas dessas propostas tenham conduzido à criação de corpora anotados, a integração entre causalidade e relações temporais é, em muitos casos, realizada através da junção de camadas de anotação desenvolvidas a partir de enquadramentos teóricos distintos. Esta opção, ainda que funcional, põe em causa frequentemente a coesão interna dos esquemas, dificultando uma representação verdadeiramente integrada das relações entre eventos.

Tendo em conta que o objetivo deste trabalho é a caracterização das relações causais em narrativas jornalísticas, e considerando a ausência de uma proposta que articule de forma coerente e unificada a informação temporal e causal, propõe-se, no Capítulo 6, o desenvolvimento de um novo esquema de anotação. Este esquema visa representar de forma harmonizada a informação relevante para a análise das relações causais, assentando num conjunto de princípios comuns que assegurem a consistência e a integração entre as duas dimensões.

6. Estudo empírico das relações causais em narrativas jornalísticas

Neste capítulo apresenta-se o estudo empírico sobre as relações causais em notícias de natureza narrativa em português europeu. De acordo com Costa et al. (2021, p. 7), qualquer “investigação visa a produção de ciência”, entendida como um mecanismo ou percurso que o ser humano, movido pela necessidade de conhecer e explicar as realidades que o rodeiam, concebeu para construir conhecimento de forma sistemática. Foi com este objetivo geral que foi desenhado um estudo que permitisse investigar as diferentes formas de produção do discurso de causalidade em notícias em português europeu, atendendo às suas estruturas, à sua organização e, sobretudo, aos elementos linguísticos que contribuem para a inferência desses valores causais. Para atingir esse objetivo geral, delineamos os seguintes objetivos específicos: identificar as relações causais em notícias em português europeu; integrar formas de anotação da causalidade no esquema de anotação do Text2Story; analisar os elementos linguísticos envolvidos na inferência das relações causais em notícias em português europeu; identificar as estruturas linguísticas mais frequentes nessas relações; identificar os dois principais eventos envolvidos; caracterizar esses eventos nas relações causais; e estabelecer as relações temporais entre os eventos de razão e de resultado.

O capítulo está organizado em duas partes. Numa primeira parte, apresentam-se os objetivos, as questões de investigação, os procedimentos metodológicos, a constituição e descrição do corpus, o esquema de anotação e o processo de anotação. Na segunda, fazem-se a apresentação e discussão dos resultados, no que diz respeito, por um lado, às relações causais (implícitas e explícitas), às estruturas que veiculam valor causal e aos sinais linguísticos e, por outro lado, às propriedades dos eventos nas relações de causa, bem como às relações temporais entre dois argumentos de causa. Por último, faz-se uma sistematização dos resultados descritos anteriormente.

6.1. Descrição do estudo

A descrição do estudo que se apresenta neste capítulo está organizada em seis secções: (i) definição dos objetivos e das questões de investigação; (ii) procedimentos

metodológicos; (iii) constituição e descrição do corpus; (iv) definição do esquema de anotação; (v) processo de anotação; e (vi) validação da anotação.

6.1.1 Objetivos e questões de investigação

Em termos gerais, este estudo tem como objetivo analisar as relações de causalidade em notícias da Agência Lusa, em português europeu, publicadas entre setembro e novembro de 2023, configurando-se como um estudo sincrónico da língua. Para alcançar este objetivo global, procede-se à caracterização das relações de causa nas notícias da Agência Lusa em português europeu, à classificação dos elementos linguísticos envolvidos na inferência das relações causais e dos tipos de sinais de causa mais frequentes no corpus, à caracterização das propriedades dos eventos que constituem os argumentos das relações causais e à correlação entre as relações temporais e as relações causais identificadas no corpus.

Neste contexto, procura-se responder às seguintes questões de investigação (QI):

QI1: Quais são as principais propriedades semânticas das relações de causa em notícias de natureza narrativa em português europeu?

QI2: Qual a configuração da organização dos argumentos das relações causais em notícias de natureza narrativa em português europeu?

QI3: Com que frequência ocorrem as relações causais explícitas e implícitas em notícias de agência em português europeu?

QI4: Quais são os elementos linguísticos que contribuem para a identificação das relações de causa em notícias de natureza narrativa em português europeu?

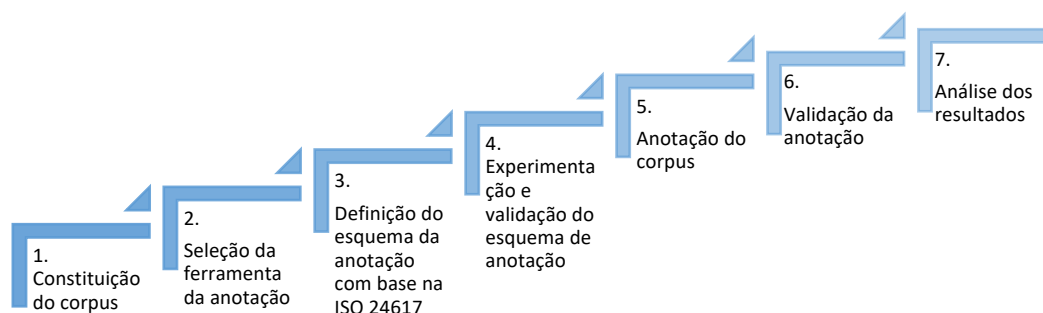
QI5: Entre as propriedades anotadas dos eventos representados nos argumentos da relação causal, que combinações de atributos ocorrem com maior frequência em notícias de natureza narrativa em português europeu?

QI6: Quais são as relações temporais estabelecidas entre os eventos principais dos argumentos nas relações causais em notícias de natureza narrativa em português europeu?

6.1.2 Procedimentos metodológicos

O procedimento metodológico desenvolveu-se de forma faseada, como se pode ver na seguinte Figura.

Figura 11 - Procedimentos metodológicos



Numa primeira etapa, procedeu-se à constituição do corpus, a partir da recolha de 1134 notícias. Seguidamente, foi feita a seleção do corpus, definindo critérios de inclusão e de descrição, nomeadamente a extensão das notícias, entre 200 e 400 palavras, e a sua temática, centrada em fenómenos naturais, tráfico e golpes de Estado. Da aplicação destes critérios resultou a seleção de 300 notícias, que integram o corpus desta investigação.

Numa segunda etapa, procedeu-se à seleção da ferramenta de anotação. Posteriormente, definiu-se o esquema de anotação causal, com base na Norma ISO 24617. Este esquema foi, numa fase seguinte, sujeito a experimentação e validação.

Após a validação do esquema de anotação, realizou-se a anotação do corpus, identificando-se um total de 829 relações causais nas 300 notícias analisadas. Seguidamente, procedeu-se à validação das anotações efetuadas.

Por fim, os resultados da anotação foram objeto de uma análise quantitativa e qualitativa, etapa que permitiu interpretar e sistematizar as informações recolhidas ao longo do processo de investigação.

6.1.3 Constituição e descrição do corpus

O tipo de texto escolhido para ser objeto de análise neste trabalho de investigação foi a notícia. Esta escolha justifica-se pelo facto de este tipo de textos favorecer a explicitação de relações de causalidade entre eventos. As notícias narrativas descrevem situações, consequências, motivações e explicações, o que conduz à ocorrência explícita ou

implícita de relações de causa e efeito. Do ponto de vista metodológico, a escolha de notícias narrativas assegura ainda uma maior homogeneidade da estrutura do corpus, uma vez que este tipo de textos obedece a convenções jornalísticas relativamente estáveis. Esta regularidade contribui para a comparabilidade dos dados e para a consistência do processo de anotação das relações causais.

No que diz respeito à estrutura das notícias, de acordo com Gonçalves (2019), recorre-se a um método de hierarquização da informação, designadamente à técnica da pirâmide invertida, na qual os factos são organizados em função da sua relevância relativa. Tipicamente, a notícia procura responder a seis questões fundamentais: *Quem?* *O quê?* *Onde?* *Quando?* *Como?* e *Porquê?*. Segundo o autor, os jornalistas estruturam as respostas a estas questões, salientando a informação mais relevante no parágrafo inicial, designado por *lead*. Nesse parágrafo, apresentam-se os participantes (*Quem?*), a ação ou o acontecimento (*O quê?*), o local (*Onde?*) e o momento (*Quando?*). No corpo do texto, são incluídas informações adicionais ou complementares que respondem às questões *Como?* e *Porquê?*, com o objetivo de fornecer contexto, explicações e detalhes que contribuam para uma melhor compreensão dos factos relatados.

Na produção das notícias da Agência Lusa, que constituem o nosso corpus, os jornalistas seguem um conjunto de normas e procedimentos que visam assegurar a uniformidade e a avaliação da qualidade editorial. Estas orientações encontram-se sistematizadas no manual intitulado *Livro de Estilo* (2019), no qual se define a estrutura padrão das notícias, assente na pirâmide invertida. De acordo com o *Livro de Estilo da Agência Lusa*, atualizado em 2019, o título deve destacar palavras-chave relevantes da notícia e não ultrapassar os 80 caracteres. O *lead* deve consistir num único período que contenha as informações essenciais relativas às questões *Quem?*, *O quê?*, *Quando?* e *Onde?*. No corpo da notícia, apresentam-se as informações adicionais ou complementares, correspondentes às questões *Como?* e *Porquê?*. Cada frase constitui um parágrafo, devendo ainda constar a assinatura profissional do jornalista responsável. A extensão da notícia não deve exceder os 3000 caracteres. As notícias, sejam de âmbito nacional ou internacional, incluem o nome da localidade e a data de publicação. Quanto à data, devem ser indicados o dia, o mês e o ano de divulgação da notícia. Relativamente

às referências temporais, utilizam-se expressões como *hoje*, *ontem*, *sexta-feira* ou *o fim de semana*.

De acordo com o *Livro de Estilo* já referido, nos casos em que são relatados acontecimentos relacionados com o intervalo de enunciação, com factos de valor real ou com a possibilidade de ocorrência, recorre-se, geralmente, ao presente do indicativo. Para a descrição de eventos já ocorridos, observa-se o uso do pretérito perfeito ou do pretérito imperfeito. Os acontecimentos são apresentados de forma cronológica, um após o outro, salientando-se os seus detalhes específicos. Procura-se enumerar a sucessão de ocorrências num determinado acontecimento com duração limitada no tempo, acompanhando os eventos hora a hora ou mesmo minuto a minuto. Um aspeto central da notícia é a sua linguagem, sendo fundamental “simplificar a compreensão das situações, devendo procurar hierarquizar a relevância das informações e preservar o sentido do que foi anunciado (e não refletir a ordem cronológica da apresentação dessas informações)” (Lusa, 2019, p. 25).

O nosso estudo empírico baseia-se num conjunto de notícias de Agência da Lusa. No que se refere à constituição do corpus desta investigação, foi efetuado um pedido formal à Agência Lusa para a recolha dessas notícias, que obteve deferimento. Recebeu-se um primeiro lote composto por 563 notícias. Dado que se pretendia que as notícias tivessem uma natureza narrativa por favorecer relações de causalidade, este primeiro as notícias deste primeiro lote foram distribuídas por duas categorias: notícias adequadas, com características narrativas e notícias não adequadas, sem características.

Após esta fase, procedeu-se à extração e ao registo das palavras-chave das notícias selecionadas. Definiram-se, então, três critérios para a seleção das notícias consideradas apropriadas para análise: a extensão das notícias, compreendida entre 200 e 400 palavras; a presença de palavras-chave associadas a fenómenos naturais, guerras, golpes de Estado, tráfico, entre outros; e o carácter narrativo dos textos.

No final deste processo, numa primeira etapa, foram selecionadas 85 notícias do primeiro lote. Contudo, identificaram-se 10 casos de repetição, o que conduziu à redução do conjunto para 75 notícias. Num segundo lote, composto por 571 notícias disponibilizadas pela mesma agência noticiosa, aplicaram-se os mesmos procedimentos

e critérios definidos anteriormente, repetindo-se o processo de seleção adotado para o primeiro lote. No final desta etapa, foram selecionadas 279 notícias, que, somadas às 75 do primeiro lote, perfizeram um total de 354 notícias.

Para o presente estudo, foram selecionadas as primeiras 300 notícias, de entre as 354 recolhidas, redigidas em português europeu. Este conjunto de notícias, de natureza narrativa, relata acontecimentos sob a forma de uma sequência de eventos, configurando uma história. As notícias foram publicadas entre setembro e novembro de 2023, abrangendo contextos nacionais e internacionais, e apresentam uma extensão entre 200 e 400 palavras. Os temas abordados são diversos, incluindo fenómenos naturais, tráfico e golpes de Estado.

Cada uma das 300 notícias possui um identificador específico, como, por exemplo, lusa2_1.txt, lusa2_10.txt, lusa2_100.txt ou lusa2_101.txt. Posteriormente, as notícias selecionadas foram reorganizadas segundo um formato específico. Na primeira linha, indica-se o local e a data de publicação. Na linha seguinte, apresentam-se os nomes abreviados dos autores. Seguem-se o título, o *lead* e, por fim, o corpo do texto. A título ilustrativo, apresenta-se de seguida um exemplo dessa organização:

Roma, 04 dez 2023 (Lusa)

JAZF/SCA // SCA

Joalheiro condenado a 17 anos de prisão por matar dois assaltantes em Itália

Um tribunal da cidade de Asti, no norte de Itália, condenou hoje a 17 anos de prisão um joalheiro acusado de ter matado dois homens e ferido um terceiro enquanto estes assaltavam a sua loja em 2021.

Mario Roggero foi declarado culpado dos homicídios de Giuseppe Mazzarino, 58 anos, e Andrea Spinelli, 44 anos, e da tentativa de homicídio de Alessandro Modica.

Apesar de ter sofrido um ferimento na perna durante o assalto à loja, situada na localidade de La Morra, na província de Cuneo, Alessandro Modica conseguiu fugir, mas foi detido pelas autoridades algumas horas após o crime.

O Ministério Público solicitou anteriormente uma pena de 14 anos para Roggero, de 68 anos.

“A palavra ‘defesa’ entra em conflito com o vídeo (do tiroteio), no qual se assiste a uma execução”, disse o procurador quando apresentou o pedido de condenação.

O vice-primeiro-ministro e ministro dos Transportes e Infraestruturas italiano, Matteo Salvini, manifestou a sua solidariedade para com Mario Roggero.

“O joalheiro defendeu a sua vida e o seu negócio”, disse Salvini, também líder do partido italiano Liga (extrema-direita).

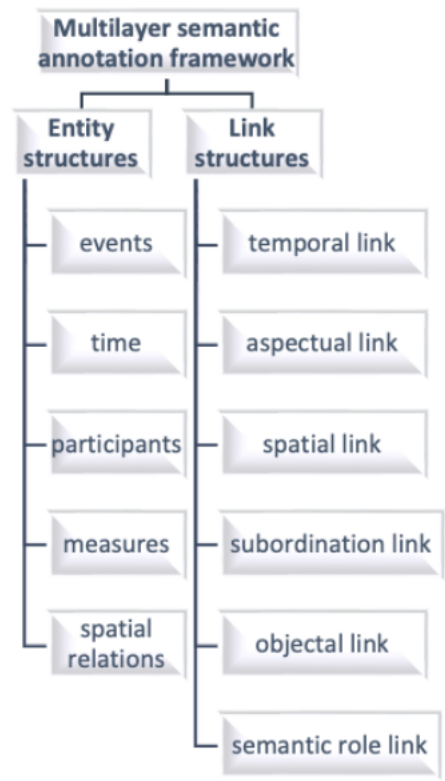
“Os verdadeiros criminosos são aqueles que merecem a prisão”, acrescentou.

6.1.4 Esquema de anotação

Neste trabalho de investigação, propomos um esquema de anotação integrado das relações causais e temporais. Partimos do esquema Text2Story¹ (cf. Figura 12), proposto por Silvano et al. (2021) e Leal et al. (2022), que tem como base a Norma 24617, pelas seguintes razões: o esquema Text2Story tem produzido resultados positivos na anotação de narrativas jornalísticas (Silvano et al. 2024) e a Norma 24617 inclui uma parte dedicada às relações discursivas (parte 8). Sendo o nosso objetivo desenvolver um esquema que seja capaz de articular de forma coerente e harmoniosa as relações causais e a informação temporal, consideramos que a melhor estratégia para o fazer é harmonizar a Parte 1 (ISO 24617-1, 2012) da Norma ISO 24617, já adaptada para o português europeu por Silvano et al. (2021) e Leal et al. (2022), com a Parte 8 (ISO 24617-8, 2016) focada nas relações discursivas (cf. Figura 13).

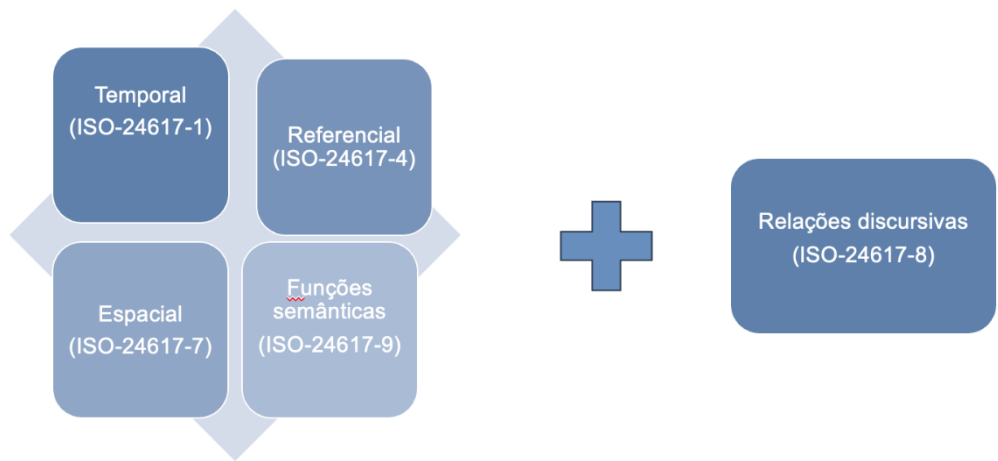
¹ Este esquema foi desenvolvido no âmbito do projeto Text2Story com a referência PTDC/CCI-COM/31857/2017 (NORTE-01-0145-FEDER-03185) (DOI [10.54499/PTDC/CCI-COM/31857/2017](https://doi.org/10.54499/PTDC/CCI-COM/31857/2017)). Para mais informações sobre o projeto, veja-se <https://text2story.inesctec.pt>.

Figura 12 - Text2Story multilayer annotation scheme.



Fonte: (Leal et al., 2022)

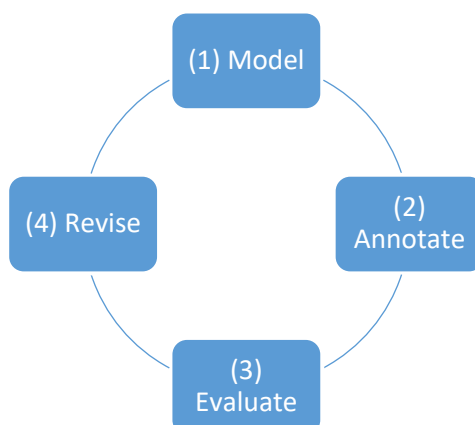
Figura 13 - Integração da Relações discursivas de causa



Quanto ao método de construção do esquema de anotação causal, seguiu-se o subciclo do modelo MATTER, proposto por Pustejovsky & Stubbs (2012), baseado num processo iterativo de *Model* e *Annotate*. Neste processo, a definição conceptual do

esquema é continuamente refinada a partir da anotação de dados reais. Posteriormente, estas etapas foram sistematizadas e explicitadas como quatro fases principais: (1) *model* (modelar), (2) *annotate* (anotar), (3) *evaluate* (avaliar) e (4) *revise* (rever) (cf. Figura 14), refletindo o caráter cíclico e incremental do desenvolvimento de esquemas de anotação.

Figura 14 - Metodologia de desenvolvimento do esquema de anotação (*MATTER cycle*)



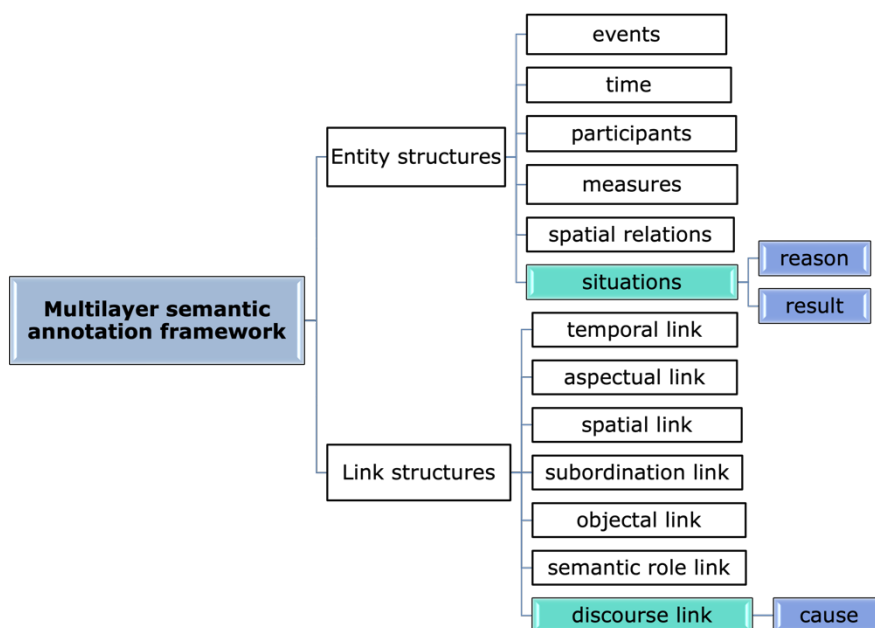
Fonte: (Pustejovsky & Stubs, 2012)

Efetivamente, este ciclo teve de ser aplicado de forma iterativa ao longo do processo de desenvolvimento do esquema de anotação uma vez que surgiram problemas de inconsistência, lacunas e incompatibilidades que exigiram o refinamento contínuo do esquema, com o objetivo de representar adequadamente os dados relacionados com as relações causais nas notícias. Como refere Silvano et al. (2021), “building a bootstrapping annotation scheme is a very complex and time-consuming endeavor involving different phases” (Silvano et al., 2021, p. 3).

A integração da Parte 8 no esquema Text2Story (Silvano et al. 2021; Leal et al. 2022) de modo a ser compatível essencialmente com a camada de anotação temporal implicou não só a projeção de mais etiquetas e atributos, como a definição de mais princípios metodológicos de anotação. A Figura 15² sistematiza a primeira versão do esquema de anotação desenvolvido.

² Optou-se por manter as etiquetas e atributos do esquema de anotação em inglês, tal como usados no esquema Text2Story e na ISO 24617. A plataforma de anotação integra as etiquetas e atributos em inglês. Por isso, os exemplos que iremos apresentar retirados desta plataforma estarão com anotações em inglês. Ao longo deste capítulo, no corpo do texto usaremos sobretudo as traduções em português.

Figura 15 - Esquema de anotação causal preliminar



Numa fase inicial foi construída, portanto, a camada causal. Para isso, acrescentou-se uma nova etiqueta, designada *Situation*, à estrutura de entidades do esquema de Silvano et al. (2021) e de Leal et al. (2022). Esta etiqueta permite anotar situações expressas sob a forma de frase, oração, nome, sintagma nominal ou verbo. A *Situation* apresenta dois papéis: razão (*reason*) e resultado (*result*), que possibilitam a anotação explícita de dois segmentos textuais como argumentos da relação discursiva Causa (*Cause*). Para estabelecer a relação causal entre estes dois segmentos, introduziu-se, nas estruturas de ligação, uma nova relação denominada relação discursiva (*Discourse link*).

Sendo um dos objetivos analisar os tipos de marcadores das relações causais, acrescentou-se a etiqueta Sinal (*Signal*), posicionada após a etiqueta Situação (*Situation*). Com base no PDTB e em Soared da Silva (1999), os sinais causais explícitos foram categorizados em seis classes: Nome, Verbo, Conjunção, Adjetivo, Advérbio e Preposição. Paralelamente, foi introduzida, nas estruturas de ligação do esquema, uma nova ligação denominada Sinal causal (*Signal causal*), que estabelece a ligação entre o Sinal e o argumento de resultado. Deste modo, conseguimos que cada sinal ficasse ligado a cada uma das relações causais explícitas observadas.

Com o intuito de testar a adequação e a funcionalidade do esquema proposto, procedeu-se à categorização e à análise da relação de causa em algumas notícias da Agência Lusa, numa fase de experimentação realizada manualmente. Nesta etapa, identificaram-se o argumento de razão, o argumento de resultado e as pistas ou elementos linguísticos que sinalizam a relação causal, como se ilustra no exemplo seguinte:

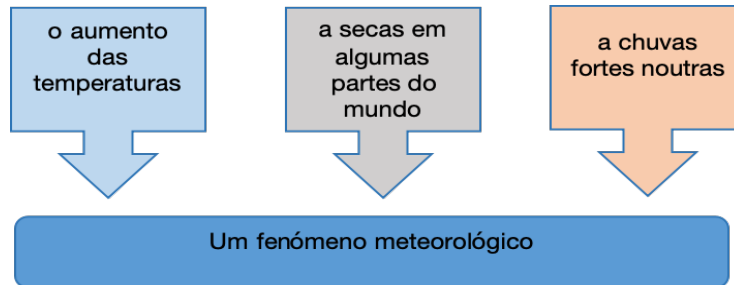
(122) O ICNF informou hoje ter determinado a “suspensão preventiva” do funcionamento do parque, assim como do bar e do restaurante do empreendimento, devido a descargas de efluentes não tratadas que põem em causa “a preservação dos valores naturais e paisagísticos do Parque Natural da Serra da Estrela (Lusa2_12).

Neste excerto, o argumento de razão encontra-se sublinhado, enquanto o argumento de resultado está marcado a itálico. O sinal causal é representado pela locução prepositiva “devido a”.

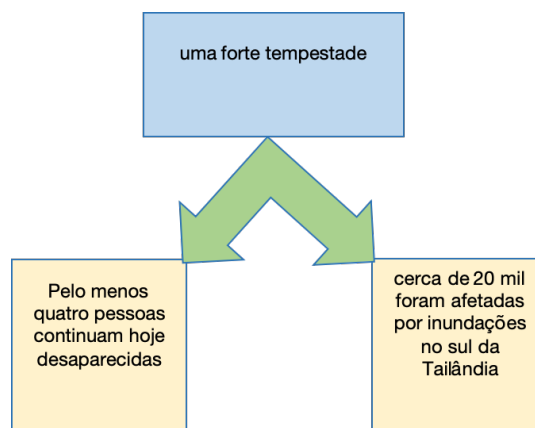
Durante a fase de experimentação do esquema de anotação, deparamo-nos com algumas questões sobre as quais tivemos de refletir e que produziram alterações no esquema e no guião de anotação, que iremos discutir a seguir.

Uma das questões prende-se com o facto de que a configuração da expressão de causalidade assume diferentes formas. O caso prototípico é aquele em que há relação casual (*Cause*) entre um argumento de razão (*reason*) e um argumento de resultado (*result*), como o ilustrado no exemplo anterior. Contudo, há outras configurações: há casos em que um único argumento de razão (sublinhado) se liga a mais do que um argumento de resultado (em itálico), como em (123); há outros casos em que várias razões podem conduzir a um único resultado, como em (124); e, noutros casos, a causa pode estar associada a alternativas, como ilustrado em (125).

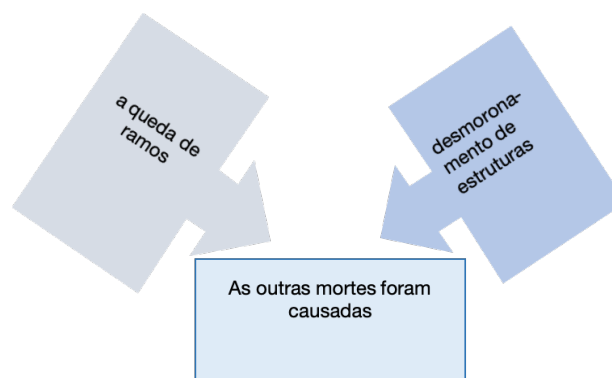
(123) Um fenómeno meteorológico geralmente associado ao aumento das temperaturas, a secas em algumas partes do mundo e a chuvas fortes noutras (Lusa2_100)



(124) Pelo menos quatro pessoas continuam hoje desaparecidas e cerca de 20 mil foram afetadas por inundações no sul da Tailândia, na sequência de uma forte tempestade (Lusa2_116)

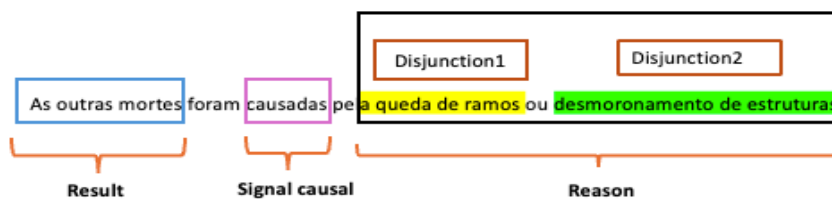
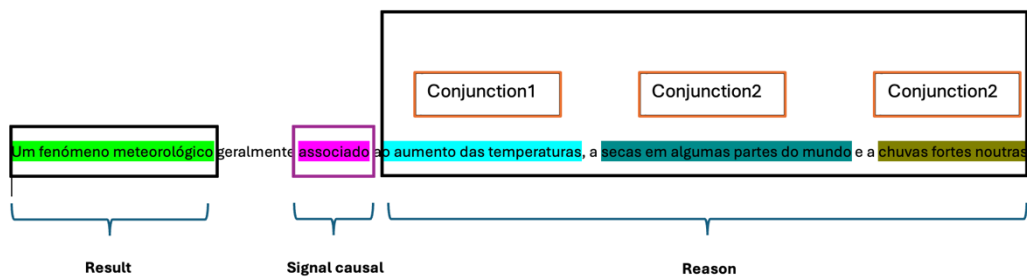


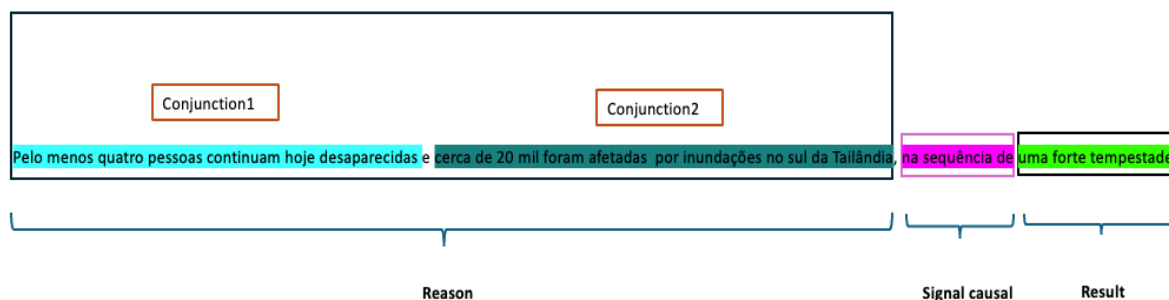
(125) As outras mortes foram causadas pela queda de ramos ou desmoronamento de estruturas (Lusa2_109)



No paradigma de anotação da Norma ISO 24617-8, a causalidade é analisada tendo em conta a relação entre dois e apenas dois argumentos, a um nível local. Para manter a coerência da anotação, em consonância com as orientações e discussões teóricas sobre relações discursivas e com a estrutura sintática dos exemplos, foram identificadas soluções adequadas para os três problemas referidos anteriormente. A proposta consiste na introdução de quatro novos papéis na etiqueta *Situation* do esquema de anotação: *Conjunction1*, *Conjunction2*, *Disjunction1* e *Disjunction2*.

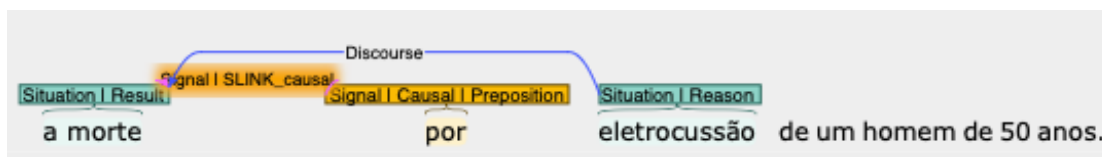
Neste enquadramento, nas três configurações que apresentamos é estabelecida apenas uma relação discursiva de causa com apenas um único argumento de razão ou de resultado. Em simultâneo, cada razão múltipla ou cada resultado múltiplo é anotado, de acordo com a ordem textual, como *Conjunction1*, *Conjunction2*, *Disjunction1* ou *Disjunction2*. Esta solução é exemplificada nas Figuras seguintes.





Outro problema de natureza técnica prende-se com uma limitação da ferramenta INCEpTion, nomeadamente a impossibilidade de adicionar constituintes pertencentes ao mesmo *markable*, isto é, segmento textual ao qual se aplica uma dada etiqueta, o que pode conduzir a falhas na anotação. O exemplo seguinte, acompanhado da respetiva anotação, ilustra esta circunstância.

(126) a morte por eletrocussão de um homem de 50 anos. (Lusa2_109)

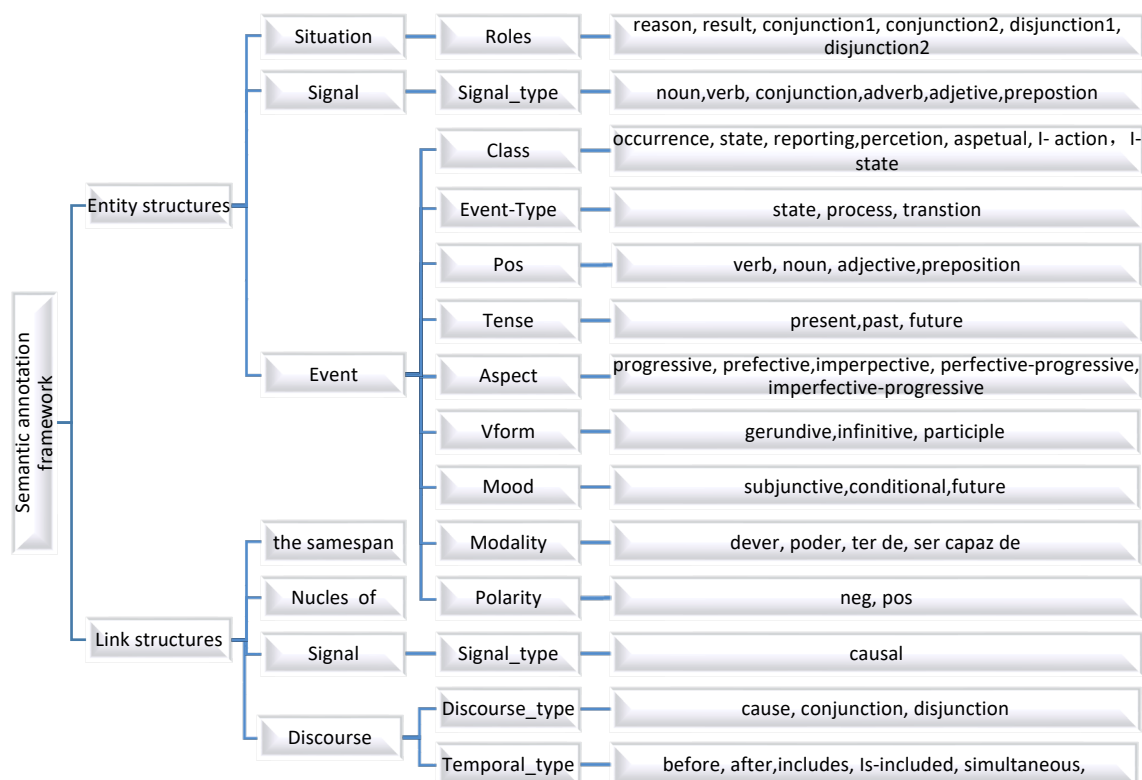


este caso, o argumento de resultado corresponde à expressão “a morte de um homem de 50 anos”. A exclusão do segmento de “um homem de 50 anos” comprometeria o significado completo da relação causal. Para preservar a integridade semântica da relação e resolver o problema de forma adequada, decidiu-se introduzir uma nova relação na estrutura de ligações do esquema preliminar, designada *sameSpan*. Assim, procede-se à anotação de “a morte” como resultado e de “um homem de 50 anos” igualmente como resultado. Posteriormente, o segundo resultado é ligado ao primeiro através da relação *sameSpan*, garantindo a continuidade da anotação do mesmo *markable* numa relação causal.

Um último problema identificado prende-se com o facto de os dois eventos principais nas relações causais se encontrarem anotados de forma autónoma, sem ligação aos argumentos dos quais fazem parte e sem qualquer ligação ao *Discourse Link*. Em relações causais simples, identificadas num único enunciado, a identificação desses eventos não levanta dificuldades significativas. No entanto, em relações causais complexas, nas quais ocorre a inserção de uma relação causal no interior de outra, torna-se difícil determinar com clareza quais os eventos que devem ser associados a cada uma das relações causais correspondentes. Para resolver esta questão, optou-se por acrescentar uma nova ligação (*link*), denominada *Nucleus of*, que estabelece a ligação entre a etiqueta *Evento* (Event) e a *etiqueta Situação* (Situation), seguindo uma direcionalidade que vai dos eventos para as situações em cada relação causal.

As fases de experimentação permitiram o aperfeiçoamento e o refinamento do esquema de anotação causal, no qual se encontram definidos as etiquetas e atributos para as entidades e relações utilizadas ao longo desta investigação durante o processo de anotação, conforme ilustrado no esquema apresentado de seguida.

Figura 16 - Esquema Lusa de Anotação Causal-Temporal (LUSA-ACT)



Na Figura, é possível observar que o esquema de anotação causal contém duas estruturas principais: as estruturas de entidades e as estruturas de ligações. A primeira inclui as etiquetas *Situation*, *Signal* e *Event*, cada uma com os respetivos atributos e valores. A segunda compreende as estruturas de ligação com as *relações the sameSpan*, *Nucleus of*, *Signal* e *Discourse*. Seguidamente, descrevem-se os elementos que integram o esquema de anotação.

No conjunto das estruturas de entidade, a etiqueta *Situation* corresponde aos papéis das relações discursivas e é realizada por uma expressão linguística simples ou complexa, como uma oração, uma nominalização, uma frase, ou um conjunto de frases. Os papéis associados à *Situation* são os seguintes: *Reason*, *Result*; *Conjunction1*, *Conjunction2*; *Disjunction1*, *Disjunction2*.

A etiqueta *Signal* refere-se a uma palavra ou a uma expressão constituída por vários vocábulos que marca explicitamente uma relação causal. Na anotação, o *Signal* não integra qualquer argumento da relação causal, limitando-se a estabelecer a ligação entre dois argumentos. Os atributos desta etiqueta correspondem aos tipos de *Signal* causal. São eles os seguintes:

- *Verb*: incorpora no seu significado a noção de causalidade, podendo ser parafraseado como *causar uma ação* ou *estado*, como em *provocar*, *causar*, *deixar*, *ficar*, *tornar*;
- *Noun*: refere-se a nomes que indicam ou evidenciam a razão ou a origem de um acontecimento, como *causa*, *origem* ou *motivo*;
- *Conjuncts*: incluem conjunções e locuções conjuncionais que ligam duas orações ou dois termos semelhantes na mesma frase para exprimir razão, motivação, causa ou explicação, tais como *porque*, *uma vez que*, *visto que*, *dado que*, *como*, *pois*;
- *Preposition*: incluem preposições ou locuções prepositivas que correspondem a palavras invariáveis que ligam dois termos da frase, explicando o significado do primeiro pelo segundo ou introduzindo uma razão, como *por*, *com*, *de*, *devido a*, *na sequência de*, *graças a*, *por causa de*;

- *Adverb*: palavras ou expressões que constroem uma relação causal, como *assim*, *consequentemente* ou *então*.

Considerando que o presente estudo também investiga as relações temporais no âmbito das relações causais, identificam-se as principais eventualidades associadas a cada argumento de uma relação de causa, estabelecendo-se, posteriormente, as relações temporais entre essas eventualidades, as quais podem ser representadas como estados ou eventos. Nesta camada temporal, a etiqueta *Event* refere-se àquilo que acontece ou ocorre, bem como a estados ou circunstâncias, sendo caracterizada por um conjunto de atributos, definidos a seguir, de acordo com o manual de anotação de Silvano et al. (2023):

- *Class*: categoriza os eventos como *occurrence* (eventualidade de natureza dinâmica. Ex. *inundar*, *matar*), *state* (eventualidade de natureza estativa. Ex. *morar*, *estar*), *reporting* (eventualidade de uma pessoa/ organização a relatar, afirmar,... Ex. *explicar*, *dizer*), *perception* (eventualidade de percepção física. Ex. *ouvir*, *ver*), *aspectual* (eventualidade que foca uma parte da eventualidade. Ex. *começar*, *acabar*), *I_State* (estado que introduz outras eventualidades como argumentos num contexto intensional. Ex. *acreditar*, *imaginar*) ou *I_Action* (evento que introduz outras eventualidades como argumentos num contexto intensional. Ex. *pedir*, *prometer*);
- *Type*: especifica os tipos aspetuais lexicais dos eventos, como *state*, *process* ou *transition* (culminação ou processo culminado);
- *Pos*: identifica as categorias gramaticais, tais como *verb*, *noun*, *adjective* ou *preposition*;
- *Tense*: representa o intervalo de tempo no qual a situação se localiza (*present*, *past*, *future*);
- *Aspect*: define o valor aspetual do evento, como *progressive*, *perfective*, *imperfective* ou *imperfective_progressive*;
- *Vform*: identifica as formas verbais não finitas, como *gerundive*, *infinitive* ou *participle*;
- *Mood*: identifica o modo verbal, como *subjunctive*, *conditional* ou *future*;

- *Modality*: representa o valor modal do evento, como *dever*, *poder* ou *ter_de*;
- *Polarity*: define a polaridade do evento, como *negative* ou *positive*.

No caso das eventualidades representadas por nomes (nominalizações ou nomes não derivados de verbos) o tipo aspetual não é anotado, por se tratar de uma classificação pouco trabalhada, não afetando, contudo, a anotação da causalidade. Para além disso, nos casos em que ocorrem adjetivos a modificar eventos, seguimos a proposta do esquema de anotação Text2Story e incluímos os adjetivos no markable dos eventos, como em *chuvas torrenciais* ou *inundações gigantescas*. As expressões lexicalizadas são igualmente consideradas como um único *markable*.

No conjunto das estruturas de ligação, há uma etiqueta *Discourse* (Discurso) que abrange o *Discourse Link Cause* (Relação Discursiva de Causa) e os *Temporal Links* (Relações Temporais). A relação causal estabelece-se quando o Arg.2 funciona como explicação do Arg.1, que assumem, assim, os papéis de Razão e Resultado. A direcionalidade da relação causal é do argumento Razão para o argumento Resultado. A relação temporal estabelece diferentes tipos de relações temporais entre duas eventualidades (*Event*) envolvidas numa relação causal. A direcionalidade das relações temporais é da eventualidade mais recente para a anterior, de acordo com a ordem linear do discurso. Os valores das relações temporais são os seguintes: *Before* - um evento ocorre antes de outro; *After* - um evento ocorre depois de outro; *Simultaneous* - um evento ocorre ao mesmo tempo que outro; *Includes* - um evento inclui outro; *Is_Included* - um evento está incluído noutro.

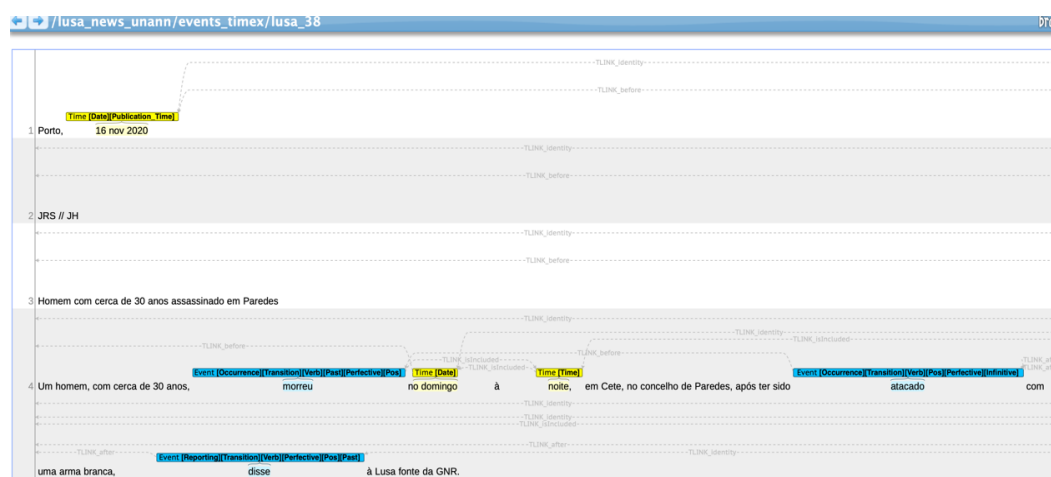
Para além destas estruturas de ligação que representam informação semântica, criamos mais três estruturas de natureza mais técnica. Assim, a ligação *Signal* associa um *Signal* causal a uma relação causal ligando-o sempre ao argumento do Resultado. A direcionalidade desta relação é do *Signal* para o argumento de Resultado. A ligação *Nucleus of* estabelece-se entre eventos e situações das relações causais, ou seja, como os argumentos de Razão e de Resultado, com direcionalidade dos eventos para as situações correspondentes. A ligação *the sameSpan* liga segmentos não contíguos pertencentes ao mesmo argumento de razão ou de resultado.

6.1.5 Processo de anotação

Antes de iniciar o processo de anotação, foi necessário escolher a ferramenta de anotação. Foram consideradas duas plataformas: o *Rapid Annotation Tool* (Brat) (Stenetorp et al., 2012) e o *INCEpTion* (Klie et al., 2018). Ambas permitem criar, editar e gerir anotações em textos e têm sido usadas em projetos de anotação sobretudo no âmbito do *Processamento de Linguagem Natural* (PLN).

O *Rapid Annotation Tool* tem sido utilizado em vários projetos, nomeadamente no projeto Text2Story, constituindo uma das ferramentas relevantes para a análise semântica. Esta plataforma é particularmente útil para copiar anotações de elementos linguísticos, promovendo a eficiência do processo de anotação. Permite ainda adicionar constituintes contínuos ao mesmo *span* e restabelecer relações entre diferentes elementos. Contudo, apresenta algumas limitações: não permite realizar curadoria entre anotações de diferentes anotadoras; deixou de receber atualizações; tende a tornar-se mais lenta quando se trabalha com textos extensos; e, ao apagar a anotação de uma entidade, é necessário eliminar previamente as anotações das relações associadas a essa entidade. Estas limitações acabam por dificultar o processo de anotação dos dados. A seguir, apresenta-se uma das interfaces desta ferramenta.

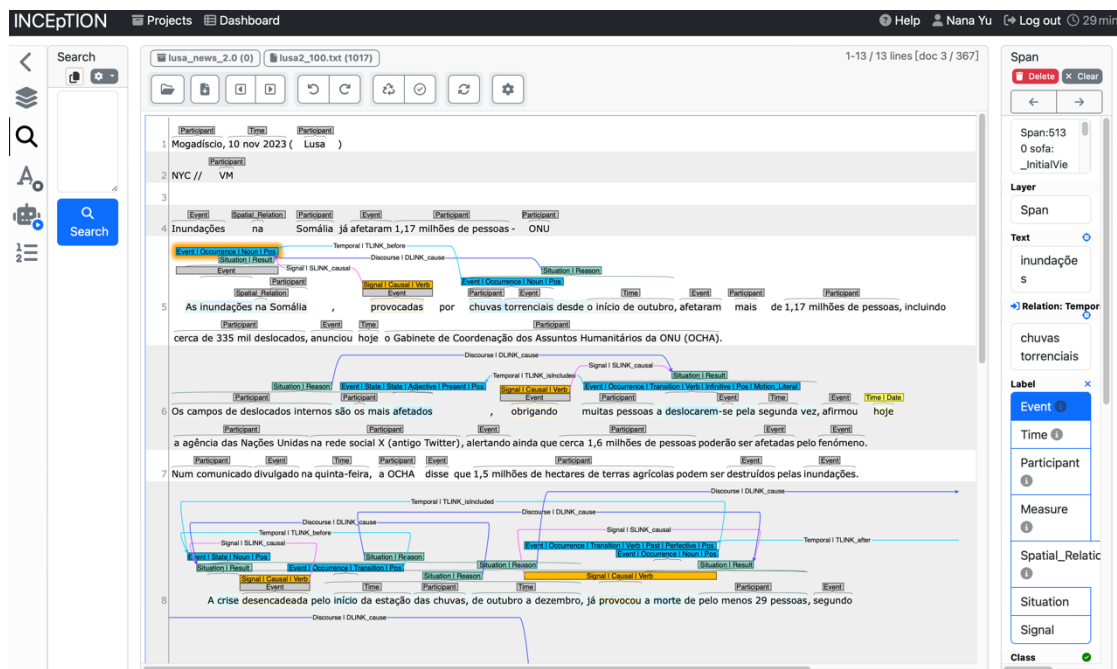
Figura 17 - Interface da plataforma Brat



No que diz respeito ao *INCEpTion*, trata-se de uma ferramenta de anotação semântica mais recente e em constante atualização. Destaca-se por várias vantagens, nomeadamente: possibilitar a colaboração em projetos mais robustos; permitir a curadoria entre anotadores; ser possível atribuir diferentes papéis, como anotadores ou

curadores; gerir projetos de forma mais eficiente; realizar pré-anotações com recurso a modelos previamente definidos no sistema; e visualizar o esquema de anotação de diferentes níveis no painel lateral direito, como se observa na figura seguinte. Ainda assim, também apresenta limitações, tais como a impossibilidade de copiar anotações, de adicionar constituintes contínuos ao mesmo *span* e de restabelecer relações entre elementos.

Figura 18 - Interface da plataforma INCEpTION



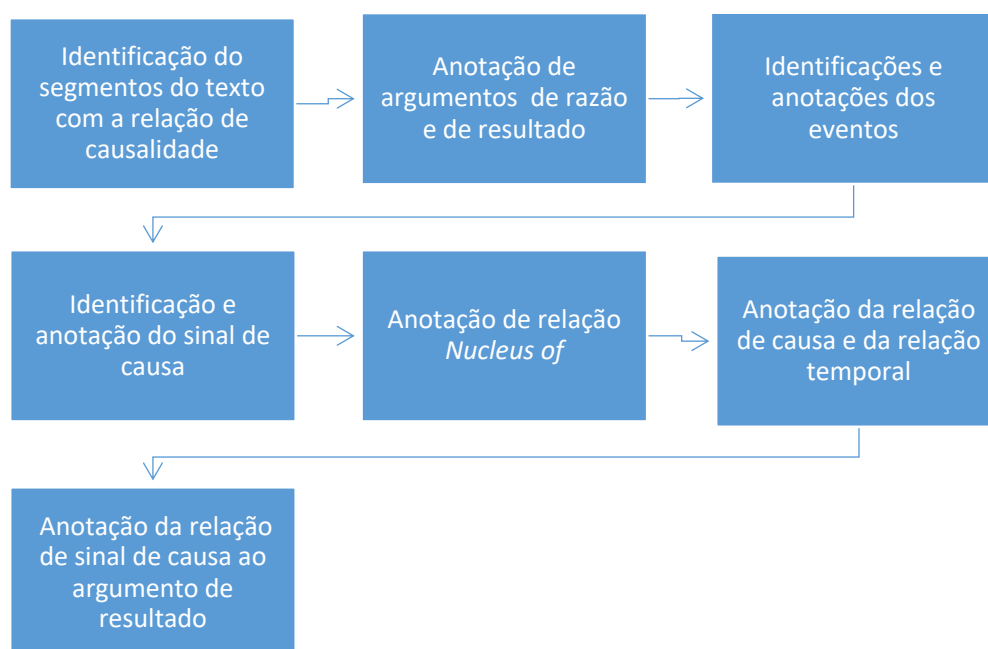
Tendo em conta a comparação entre as vantagens e desvantagens das duas ferramentas, optou-se pelo *INCEpTion* como plataforma de anotação para a análise do corpus deste estudo. Esta escolha deve-se aos pontos positivos anteriormente referidos e às experiências de utilização no âmbito do projeto *CLEF-2024 CheckThat! Lab Task 3 on Persuasion Techniques* (Piskorski et al., 2024).

Selecionada a plataforma de anotação, esta foi configurada. Como descrevemos na secção anterior, a anotação causal envolve diferentes etiquetas e relações. Para tornar as anotações mais visíveis na plataforma de anotação, antes do início do processo de anotação definiram-se cores específicas para cada tipo de etiqueta e para cada tipo de relação. Os argumentos de Razão, de Resultado e de Disjunção estão assinalados a

verde, enquanto os argumentos de Conjunção são representados a roxo. O sinal causal surge a cor-de-rosa. Os dois eventos principais de uma relação causal encontram-se anotados a azul. A relação discursiva de causa e a respetiva relação temporal são representadas por uma linha verde contínua, ao passo que a relação de sinal causal é indicada a cor-de-rosa.

De seguida, definimos as etapas a seguir durante o processo de anotação dos dados, ilustradas na Figura seguinte.

Figura 19 - Processo de anotação causal



A primeira etapa consiste em fazer uma primeira leitura da notícia para uma compreensão global do texto. Na segunda leitura, procura-se identificar um segmento textual em que esteja presente uma relação causal, como no seguinte exemplo.

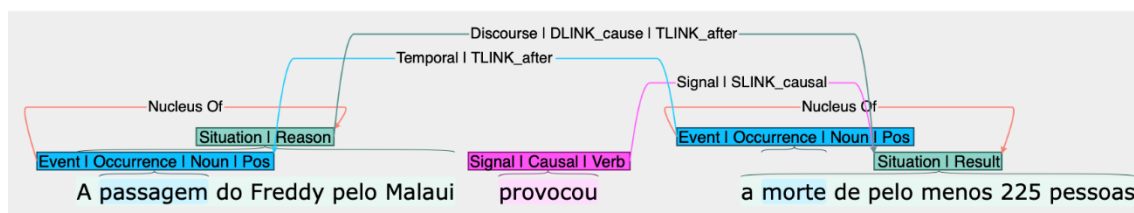
(127) A passagem do Freddy pelo Malaui provocou a morte de pelo menos 225 pessoas (Lusa2_182).

Em seguida, procede-se à identificação e anotação do argumento de Razão (*A passagem do Freddy pelo Malaui*) e do argumento de Resultado (*a morte de pelo menos 225 pessoas*). Posteriormente, identificam-se e anotam-se os dois eventos associados a

cada um dos argumentos que compõem a relação de causa (*passagem; morte*). Por último, identifica-se e anota-se o sinal de causa (*provocou*).

Relativamente à anotação das relações, começa-se por estabelecer as relações de *Nucleus of* a partir dos eventos identificados para os argumentos de Razão e de Resultado. Em seguida, anota-se a relação causal e temporal entre esses dois argumentos. Por fim, estabelece-se a relação de sinal, ligando o sinal causal ao argumento de Resultado. Deste modo, conclui-se o processo de anotação dos aspetos envolvidos numa relação de causa. O resultado final é ilustrado na seguinte Figura.

Figura 20 - Anotação de um segmento textual do corpus



Nos casos mais complexos, em que podem ocorrer mais do que um argumento de Razão e/ ou de Resultado, ou ainda razões ou resultados alternativos, é necessária uma leitura mais minuciosa e atenta aos aspetos menos explícitos do enunciado, exigindo-se a aplicação de passos adicionais com base no procedimento definido para as relações de causa simples. Segue-se um exemplo de uma relação de causa complexa.

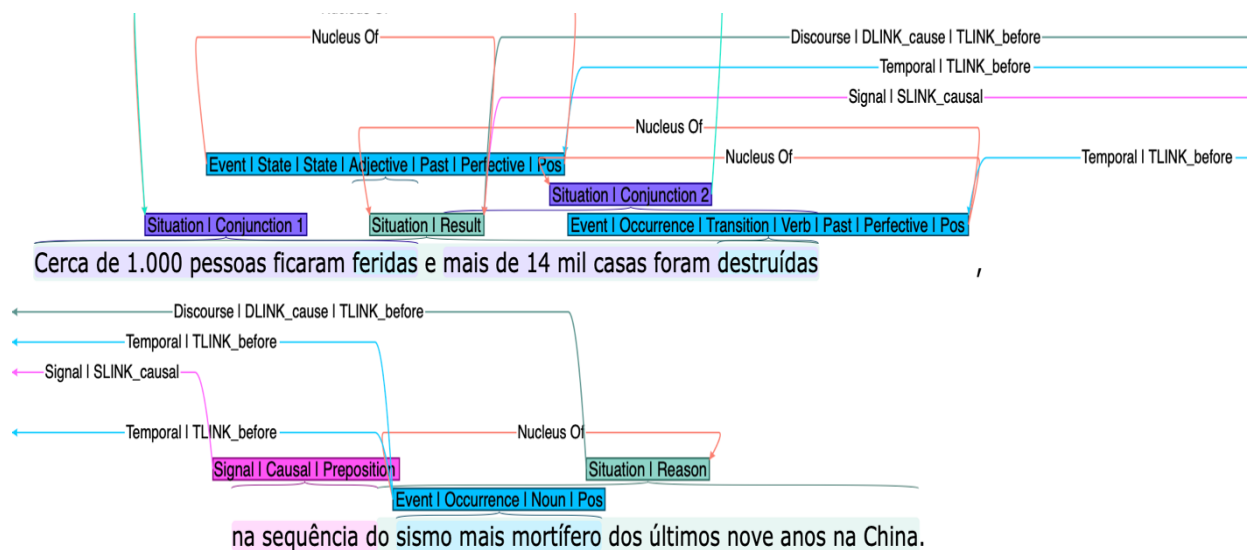
(128) Cerca de 1.000 pessoas ficaram feridas e mais de 14 mil casas foram destruídas na sequência do sismo mais mortífero dos últimos nove anos na China. (Lusa2 _129)

Em termos da anotação das entidades, começa-se por identificar e anotar as conjunções presentes no enunciado, nomeadamente a Conjunção1 (*Cerca de 1.000 pessoas ficaram feridas*) e a Conjunção2 (*mais de 14 mil casas foram destruídas*). Seguidamente, anota-se o argumento de Resultado (*Cerca de 1.000 pessoas ficaram feridas e mais de 14 mil casas foram destruídas*), constituído pela coordenação dos dois argumentos da relação discursiva Conjunção. Segue-se a anotação do argumento de

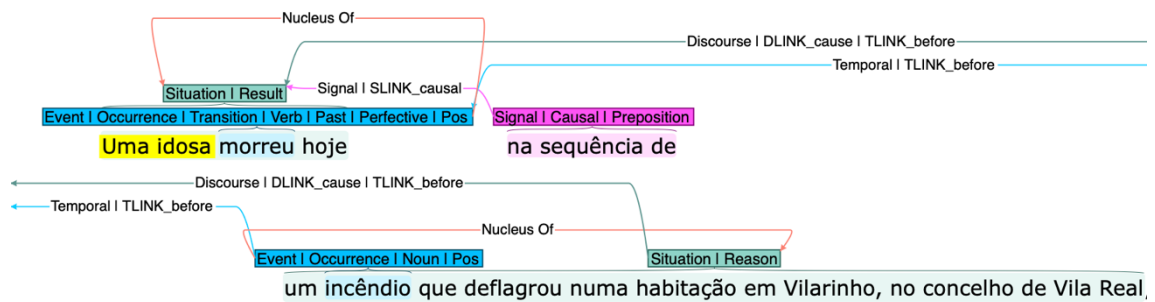
Razão (o sismo mais mortífero dos últimos nove anos na China). Posteriormente, identificam-se e anotam-se o evento (*feridas*) da Conjunção1 e o evento (*destruídas*) da Conjunção 2, assim como o evento (*sismo mais mortífero*) associado ao argumento de Razão. Por último, procede-se à anotação do sinal de causa (*na sequência de*).

Quanto à anotação das relações, estabelece-se, em primeiro lugar, a relação de *Nucleus of* entre cada evento e o respetivo argumento do par da relação de Conjunção, bem como entre esses eventos e o argumento de Resultado. De seguida, o evento associado ao argumento de Razão é igualmente ligado a esses argumentos através da relação de *Nucleus of*. Posteriormente, define-se a relação causal e temporal entre os argumentos de Razão e de Resultado. Por fim, estabelece-se a relação de sinal causal, concluindo-se assim o processo de anotação da relação de causa. A Figura seguinte mostra o resultado de todos estes passos.

Figura 21 - Anotação de um segmento textual com uma relação causal com conjunção



(129) Uma idosa morreu hoje na sequência de um incêndio que deflagrou numa habitação em Vilarinho, no concelho de Vila Real.



Após a definição do processo de anotação, procedeu-se à anotação das relações de causa em 300 notícias, com base nos conhecimentos adquiridos ao longo da revisão da literatura, nos estudos realizados e no enquadramento esquema e manual de anotação definidos para este projeto.

6.1.6 Processo de validação

De modo a garantir a qualidade e a fiabilidade dos resultados da anotação dos dados, implementaram-se alguns mecanismos de validação das anotações. Assim, todo o processo de anotação foi acompanhado pela Professora orientadora, uma especialista em Semântica, experiente em anotação. Ao longo do período de anotação, realizaram-se reuniões regulares com o objetivo de esclarecer dúvidas e resolver problemas decorrentes da anotação.

Após a conclusão da anotação de todas as notícias por parte da anotadora, o método de validação consistiu, numa primeira fase, na revisão das anotações das primeiras dez notícias pela Professora orientadora. Os problemas identificados foram registados e procedi à correção dos mesmos tipos de erros nas restantes notícias. Concluídas essas correções, a Professora orientadora fez a verificação de mais um conjunto de dez notícias. Este processo de feedback repetiu-se até deixarem de ser identificados erros, dando-se assim por concluído o processo de validação dos dados.

Durante este processo, foram identificados alguns casos que não são considerados relações de causa no âmbito deste trabalho de investigação. Incluem-se, nomeadamente, os casos em que o argumento de razão ou de resultado corresponde a um participante em vez de uma eventualidade, como em (130); os casos em que um dos argumentos se refere a um numeral (131); os casos em que apenas um dos argumentos (razão ou resultado) está presente, não se verificando a ocorrência de ambos (133); os

casos em que o significado do enunciado não se encontra suficientemente claro (132); os casos em que, apesar da presença de um sinal de causa, não se identifica uma relação causal, pois o primeiro argumento não inclui um evento (133); e, por fim, os casos em que a razão ou o resultado são introduzidos por um pronome relativo (134).

(130) milhares de desaparecidos, bairros inteiros arrasados e levados por **um mar**. (Caso com participante)

(131) **outras duas** por falta de documentos aquando da fiscalização (contraordenação leve) (Caso com numeral)

(132) Esta ação (descargas) coloca e, causa valores naturais e paisagísticos de interesse relevante, que apresentam sensibilidade ecológica, e **porque** impera que seja evitada a continuação da prática, sob pena da situação se tornar irreversível ou implicar uma maior dificuldade ou impossibilidade de reposição da situação à infração. (Caso sem clareza do sentido da frase)

(133) O número de mortos **provocados** por um deslizamento de terras na Colômbia subiu para 33. (Caso com verbo causal sem resultado)

(134) Nesse período, o subcontinente indiano recebe 70% da precipitação anual, **o que** frequentemente conduz a inundações e outras catástrofes naturais. (Caso de oração relativa)

6.2. Apresentação e discussão dos resultados

Nesta secção central do trabalho, apresentam-se, num primeiro momento, a caracterização e a descrição dos dados recolhidos, dando a conhecer a sua organização e categorização. Num segundo momento, procede-se à análise e interpretação dos dados empíricos, com o objetivo de compreender as diferentes formas de ocorrência das relações de causa nas notícias da Agência Lusa publicadas entre outubro e novembro de 2024.

Como já foi previamente referido, o corpus recolhido é composto por um total de 300 notícias redigidas em português europeu pela Agência Lusa, publicadas no

período de outubro a novembro de 2024. Os dados anotados são apresentados quantitativamente sob a forma de tabelas e esquemas, seguidos de descrição e análise.

Destacam-se, assim, as frequências máxima, média e mínima de relações de causa por notícia, tendo em conta o facto de estas poderem ser explícitas ou implícitas, bem como simples ou complexas, considerando a presença de conjunções e disjunções. Analisam-se ainda a distribuição das marcas linguísticas de causalidade (verbos, nomes, conjunções, preposições e advérbios), as propriedades semânticas dos argumentos de razão e de resultado, as relações temporais entre os eventos dos argumentos de razão e de resultado e a caracterização dos eventos causais.

As atividades de organização, análise e discussão dos resultados são orientadas pelos objetivos gerais e específicos do presente estudo, bem como pelo quadro teórico desenvolvido anteriormente, com particular destaque para a perspetiva da norma ISO 24617-8 sobre a causalidade e da ISO 24617-1 sobre as propriedades temporais.

6.2.1. As relações causais no corpus

Após a análise das 300 notícias, foram identificadas 829 relações causais. Verificou-se que uma notícia pode conter várias relações causais ou não apresentar qualquer relação desse tipo. De acordo com os resultados obtidos, o número máximo de relações causais por notícia atinge 10, a média é de 2,787 e o valor mínimo é 0, como se pode verificar na Tabela seguinte.

Tabela 13 - Frequência de relações causais por notícia

Relações causais /Notícia	
	OCORRÊNCIAS
Máx.	10
Méd.	2.763
Min.	0

O valor 0, correspondente à ausência de relações causais em determinadas notícias, parece sugerir ainda que estas notícias nem sequer têm um segmento textual que responda a uma das questões que caracterizam a estrutura das notícias (Gonçalves, 2019), neste caso, o *porquê*. Trata-se de uma questão que deveria ser abordada no

corpo da notícia na estrutura da pirâmide invertida, conforme orientações do *Livro de Estilo* (2019).

A análise permite concluir que, na maioria das notícias da Agência Lusa, predomina uma média de 2,763 relações causais por notícia. É claro que nem todas as ocorrências correspondem a uma resposta ao porquê do evento que é noticiado. De qualquer modo, este resultado reforça a importância da causalidade como um dos princípios fundamentais para a compreensão e explicitação da informação relativa aos acontecimentos. Tal perspectiva encontra-se já destacada em Couper-Kuhlen & Kortmann (2000), que defendem a causalidade como um princípio indispensável na organização da vida quotidiana, permitindo aos indivíduos compreender as razões subjacentes à ocorrência de fenómenos naturais e sociais.

Após esta análise quantitativa, procede-se à análise da estrutura, organização e ordem dos argumentos que compõem as relações causais identificadas.

6.2.1.1. Estrutura e organização das relações de causa

Na análise das relações causais identificadas no corpus, observa-se que os argumentos que constituem uma relação de causa são, em regra geral, dois, desempenhando papéis distintos e complementares na construção da noção de causalidade.

Como já referimos, neste trabalho, adotamos a terminologia da ISO 24617-8 (2016) e usamos as seguintes designações para os argumentos da relação causal: Argumento 2 (Razão) e Argumento 1 (Resultado). Ao longo dos capítulos anteriores, tivemos oportunidade de mostrar que outros autores recorrem a designações distintas, como *causa* e *efeito* (Lopes, 2004; Asher & Lascarides, 2003; Barik et al., 2016), *evento-causa* e *evento-efeito* (Arrais, 1985) ou ainda *causa* e *efeito* ou *agente* e *paciente* (Soares da Silva, 1999).

Partindo do exemplo seguinte, podemos ilustrar a análise dos exemplos em termos de argumentos.

(135) Pelo menos 19 pessoas morreram na sequência de um aluimento de terras na
província de Sichuan, no sudoeste da China (Lusa2_200)

Neste exemplo, o Arg.2 (Razão), *um aluimento de terras na província de Sichuan, no sudoeste da China*, explica a causa, isto é, o *porquê* do acontecimento expresso no Arg.1 (Resultado), *Pelo menos 19 pessoas morreram*.

Para além destas estruturas mais frequentes, identificaram-se casos que podem ser considerados complexos, nos quais um argumento apresenta múltiplas razões ou múltiplos resultados, ligados por conetores aditivos (*e*) ou disjuntivos (*ou*). Como já foi explicado, na nossa proposta, nestes casos, cada razão ou resultado é anotado como argumento da relação discursiva *Conjunção* com os papéis *Conjunção1* e *Conjunção2*, *Disjunção1* e *Disjunção2*, sendo que a combinação destes dois argumentos constituirá o argumento de razão ou de resultado.

Conforme se observa na tabela, das 829 relações causais identificadas, 118 casos, correspondentes a 19,06%, são relações complexas. Destas, as mais frequentes são as ligadas por *conjunção* aditiva (95,76%), o que evidencia uma das características da estrutura dos argumentos de razão e resultado nas notícias da Agência Lusa em português europeu. Os casos de *conjunção* encontram-se ilustrados nos enunciados (136) e (137), enquanto os de *disjunção* surgem nos enunciados (138) e (139).

Tabela 14 - Casos de relações causais com conjunções e disjunções

RELAÇÕES CAUSAIS	OCORRÊNCIA	PERCENTAGEM
Conjunção	113	95,76%
Disjunção	5	4,24%
TOTAL	118	100%

(136) Pelo menos quatro pessoas continuam hoje desaparecidas e cerca de 20 mil foram afetadas por inundações no sul da Tailândia, na sequência de uma forte tempestade (Lusa2_116).

Arg. 2 (Razão) – *uma forte tempestade*

Arg.1 (Resultado)	{	Conjunção1 – <i>Pelo menos quatro pessoas continuam hoje desaparecidas</i> Conjunção2 – <i>cerca de 20 mil foram afetadas</i>
----------------------	---	---

(137) As autoridades japonesas pediram hoje precaução redobrada devido ao temporal de chuva e neve nas zonas afetadas pelo sismo de 01 de janeiro, no qual morreram pelo menos 180 pessoas (Lusa2_123)

Arg.2 (Razão)	{	Conjunção1 – <i>o temporal de chuva</i> Conjunção2 – <i>neve nas zonas afetadas pelo sismo de 01 de janeiro, no qual morreram pelo menos 180 pessoas</i>
------------------	---	--

Arg. 1 (Resultado) – *As autoridades japonesas pediram hoje precaução redobrada*

(138) Pelo menos 11 pessoas morreram por afogamento ou queda de árvores durante as tempestades que atingiram o leste da Austrália esta semana (Lusa2_119)

Arg. 1 (Resultado) – *Pelo menos 11 pessoas morreram*

Arg.2 (Razão)	{	Disjunção1 – <i>afogamento</i> Disjunção2 – <i>queda de árvores durante as tempestades que atingiram o leste da Austrália esta semana</i>
------------------	---	---

(139) o mau tempo pode causar aluimentos de terra e avalanchas ou o desmoronamento de edifícios e casas já instáveis (Lusa2_123)

Arg. 2 (Razão) – o mau tempo

Arg.1 (Resultado)	{	Conjunção1 – <i>aluimentos de terra</i> Conjunção2 – <i>avalanchas</i>
----------------------	---	--

Arg.1	{	Disjunção1 – aluimentos de terra e avalanchas
(Resultado)		Disjunção2 – o desmoronamento de edifícios e casas já instáveis

Estes resultados, ilustrados pelos exemplos anteriores, revelam a diversidade de configurações sintáticas e semânticas das relações causais.

A nossa análise permitiu ainda observar que a relação causal pode ocorrer em diferentes níveis textuais: a nível local, entre enunciados adjacentes e a um nível mais global da estrutura do texto, correspondendo a relações interoracionais e interfrásicas (Lopes & Carrilho). De acordo com a Norma ISO 24617-8 (2016), uma relação é considerada interfrásica quando envolve relações semânticas, coesivas e de conexão que ligam frases distintas num texto, recorrendo ao uso de conjunções (*e, mas, porque*) ou de pronomes que asseguram a continuidade discursiva, como se ilustra no exemplo (140). Por outro lado, fala-se de relação intrafrásica quando a relação causal ocorre no interior de uma única frase, em que o argumento de razão e o argumento de resultado se encontram integrados numa mesma estrutura frásica, como em (141). Existem ainda situações em que a relação causal se estabelece entre duas frases distintas, como no exemplo (142).

(140) O número de mortos pode aumentar porque as equipas de socorro continuam a remover as toneladas de terra que caíram sobre os condutores e passageiros.
(Lusa2_125)

(141) A passagem do Freddy pelo Malauí provocou a morte de pelo menos 225 pessoas
(Lusa2_182)

(142) “Suspeita-se que o condutor terá tido um problema de saúde, parece que terá desmaiado. Por isso, é que terá acontecido o despiste”, em que o veículo “embateu e partiu as grades e ficou dentro da escola secundária” (Lusa2_274).

No exemplo (140), identificam-se dois argumentos: o Arg.2 (Razão) (*as equipas de socorro continuam a remover as toneladas de terra que caíram sobre os condutores*

e passageiros) e o Arg.1 (Resultado) (*O número de mortos pode aumentar*). Estes argumentos correspondem a orações distintas, ligadas por uma conjunção subordinativa adverbial causal (*porque*). Já no exemplo (141), o Arg.2 (Razão) (*A passagem do Freddy pelo Malaui*) corresponde a um sintagma nominal com função de sujeito, enquanto o Arg.1 (Resultado) (*a morte de pelo menos 225 pessoas*) integra o sintagma verbal com função de predicado. No exemplo (142), a causalidade estabelece-se entre duas frases distintas: o Arg.2 (Razão) (*Suspeita-se que o condutor terá tido um problema de saúde, parece que terá desmaiado*) e o Arg.1 (Resultado) (*é que terá acontecido o despiste*).

Observou-se igualmente que, no mesmo exemplo, um argumento pode desempenhar funções distintas, surgindo como resultado numa relação causal e como razão numa relação subsequente, como se verifica no exemplo (143).

(143) Pelo menos oito pessoas morreram devido ao mau tempo na Índia, devido à aproximação à costa leste do país do ciclone Michaung, que deve atingir nas próximas horas o estado de Andhra Pradesh (Lusa2_109).

Neste exemplo, identificam-se duas relações causais. Na primeira, o Arg.2 (Razão) (*o mau tempo na Índia*) relaciona-se com o Arg.1 (Resultado) (*Pelo menos oito pessoas morreram*). Na segunda relação causal, estabelece-se uma ligação entre o Arg.2 (Razão) (*a aproximação à costa leste do país do ciclone Michaung, que deve atingir nas próximas horas o estado de Andhra Pradesh*) e o Arg.1 (Resultado) (*o mau tempo na Índia*). Assim, o argumento *o mau tempo na Índia* desempenha simultaneamente dois papéis distintos: funciona como razão na primeira relação causal e como resultado na segunda. Outra configuração observada é a ocorrência de uma relação causal interna a um Arg.2 (Razão) ou a um Arg.1 (Resultado) de uma relação causal, como se verifica no exemplo (144).

(144) As 10 pessoas morreram em consequência da "erosão causada pelas fortes chuvas" na aldeia de Miringati, território de Lubero, província do Kivu Norte (Lusa2_192)

No exemplo (144), identifica-se, em primeiro lugar, uma relação causal em que o Arg.2 (Razão) (*a erosão causada pelas fortes chuvas na aldeia de Miringati, território de Lubero, província do Kivu Norte*) se relaciona com o Arg.1 (Resultado) (*As 10 pessoas morreram*). Observa-se ainda que, no interior do Arg.2, ocorre uma segunda relação causal, na qual o Arg.2 (Razão) (*as fortes chuvas na aldeia de Miringati, território de Lubero, província do Kivu Norte*) se relaciona com o Arg.1 (Resultado) (*a erosão*).

A análise da assimetria das relações causais presentes nas notícias da Agência Lusa em português europeu revela uma lógica de interdependência. De acordo com a Norma ISO 24617-8 (2016), a interpretação do argumento de razão depende do argumento de resultado, o que significa que a relação causal só é percebida ou inferida a partir da dependência estabelecida entre ambos os argumentos (Dunietz et al., 2017). De um modo geral, a causalidade constitui uma relação de interdependência (Arrais, 1985), na qual uma oração complementa o sentido da outra, como se observa no exemplo (145).

(145) Como as chuvas não abrandam, as autoridades regionais prolongaram o alerta vermelho até domingo. (Lusa2_197)

Por fim, verifica-se que os argumentos que compõem as relações causais, Arg.2 e Arg.1, para além de poderem assumir diferentes configurações sintáticas, não seguem sempre a mesma ordem. Em alguns exemplos, ocorre primeiro o Arg.2 (Razão) seguido do Arg.1 (Resultado), enquanto noutros se observa a ordem inversa. Esta organização dos argumentos também já tinha sido atestada por Prasad et al. (2008). A Tabela seguinte sistematiza os resultados da análise da ordem dos argumentos no corpus.

Tabela 15 - Ordem dos argumentos nas relações causais

ORDEM DOS ARGUMENTOS	OCORRÊNCIA	PERCETAGEM
Razão-Resultado	303	36,55%
Resultado-Razão	526	63,45%
TOTAL	829	100%

A Tabela anterior, relativa à ordem dos argumentos nas relações causais, mostra que, das 829 relações causais identificadas nas notícias em português europeu, em 63,45% dos casos ocorre com a ordem Resultado-Razão (Arg.1 antes de Arg.2), como ilustrado no exemplo (147). Em contrapartida, em 36% das ocorrências, surge primeiro o Arg.2 (Razão), seguido do Arg.1 (Resultado), como no exemplo (146). Apresentam-se, de seguida, dois exemplos ilustrativos.

(146) As inundações **provocaram** mais de 305 deslizamentos de terras (Lusa2_197).

(147) no mês passado, cinco membros da mesma família (um casal e os três filhos) morreram depois de um deslizamento de terras ter atingido a sua casa(Lusa2_203).

A ordem razão-resultado mais frequente pode justificar-se por duas razões: (i) o jornalista está a relatar um determinado evento já introduzido na notícia, construindo o tópico e apresentando as consequências desse evento no argumento de resultado através da relação discursiva de Causa; (ii) o jornalista pretende dar mais ênfase à razão e coloca-a logo no início da frase. Há ainda uma terceira hipótese relacionada com a ordem dos argumentos refletir a ordenação temporal dos eventos. Discutiremos esta hipótese na secção dedicada às relações temporais.

6.2.1.2. Relações causais explícitas e implícitas nas notícias

As relações causais podem manifestar-se de forma explícita ou implícita, o que significa que pode ter um marcador da relação ou a causalidade ser inferida a partir de outros elementos lexicais, da semântica composicional e do conhecimento do mundo. Como refere Girju (2003),

explicit causation patterns can contain relevant keywords such as cause, effect, consequence, but also ambiguous ones such as generate, induce, etc. The implicit causative constructions are more complex, involving inference based on semantic analysis and background knowledge. (Girju, 2003, p. 3).

Neste contexto, ao observar o corpus em análise, procurou-se determinar se as relações causais identificadas eram explícitas ou implícitas e, no caso de serem explícitas, quais as diferentes formas de marcação da causalidade. A Tabela seguinte sistematiza esta análise.

Tabela 16 - Relações Causais explícitas e Implícitas no corpus

RELAÇÕES CAUSAIS	OCORRÊNCIA	PERCENTAGEM
Explícita	661	79,73%
Implícita	168	20,27%
TOTAL	829	100%

Das 829 ocorrências de causalidade identificadas nas 300 notícias, verificou-se que predominam as relações causais explícitas, num total de 661 ocorrências (79,73%), em comparação com as relações implícitas, que totalizam 168 casos (20,27%). Apresentam-se de seguida exemplos de relações explícitas.

(148) As inundações **provocaram** mais de 305 deslizamentos de terras (Lusa2_197).

(149) Mais de 36.000 pessoas foram desalojadas **devido às** inundações no nordeste de Itália (Lusa2_197).

(150) Os relâmpagos foram a principal **causa** de morte entre as catástrofes relacionadas com o clima na Índia em 2022, com 907 mortes (Lusa2_105).

(151) O combate às chamas foi dificultado **porque** os pontos de água no local “não tinham a pressão ideal para os veículos dos bombeiros, que fazem bastante consumo”. (Lusa2_317).

(152) os constrangimentos **resultantes** das condições adversas (Lusa2_134).

(153) Após a libertação, o homem ausentou-se “para parte incerta”, tendo, posteriormente, sido localizado e detido em Espanha, e **consequentemente** extraditado para os Estados Unidos da (...) (Lusa2_143).

- (154) A doença do legionário, **provocada** pela bactéria 'Legionella pneumophila' (Lusa2-137)
- (155) O Khanun chegou a Taiwan uma semana depois de o tufão Doksuri ter atingido a ilha, **causando** pelo menos um morto e cortes de energia em milhares de casas (Lusa2_206)
- (156) Uma derrocada ocorrida esta manhã está a **causar** constrangimentos na Estrada Nacional 2 (EN2), junto ao Peso da Régua. (Lusa2_132)

Como se pode observar nos exemplos, as relações causais explícitas podem ser marcadas por um conjunto diverso de elementos lexicais. No exemplo (148), a relação de causa entre o Arg.1 (Resultado) (*mais de 305 deslizamentos de terras*) e o Arg.2 (Razão) (*As inundações*) é explícita, sendo marcada pela forma verbal *provocaram*. No exemplo (149), a relação de causalidade entre o Arg.1 (Resultado) (*Mais de 36.000 pessoas foram desalojadas*) e o Arg.2 (Razão) (*às inundações no nordeste de Itália*) é explicitada pela locução prepositiva *devido a*. No exemplo (150), a relação causal entre o Arg.1 (Resultado) (*morte entre as catástrofes relacionadas com o clima na Índia em 2022, com 907 mortes*) e o Arg.2 (Razão) (*Os relâmpagos*) é explicitada por um nome com valor causal, *causa*. No exemplo (151), a relação de causa entre o Arg.1 (Resultado) (*o combate às chamas foi dificultado*) e o Arg.2 (Razão) (*os pontos de água no local não tinham a pressão ideal para os veículos dos bombeiros*) é marcada pela conjunção causal *porque*. No exemplo (152), a relação causal entre o Arg.1 (Resultado) (*os constrangimentos*) e o Arg.2 (Razão) (*as condições adversas*) é assinalada pelo adjetivo (*resultantes*), com valor causal. No exemplo (153), a relação de causalidade entre o Arg.1 (Resultado) (*extraditado para os Estados Unidos da América*) e o Arg.2 (Razão) (*Após a libertação, o homem ausentou-se “para parte incerta”, tendo, posteriormente, sido localizado e detido em Espanha*) é marcada pelo advérbio *consequentemente* (Castilho, 2014). Para além dos casos anteriores, as relações causais explícitas podem também ser identificadas através de verbos causativos em formas não finitas, nomeadamente no particípio passado (154), no gerúndio (155) e no infinitivo (156).

No que diz respeito às relações causais implícitas, identificaram-se 168 ocorrências, correspondentes a 20,27% do total das relações analisadas. Embora não haja marcadores lexicais explícitos, outros fatores permitem inferir a relação causal. Os exemplos seguintes ilustram este tipo de relações.

(157) houve um tiroteio e morreram 13 pessoas (Lusa2_13)

(158) o morador, cuja idade não soube precisar, sofreu escoriações quando ocorreu o desabamento (Lusa2_328)

(159) "Esta é uma tragédia nacional. Apelo aos parceiros e doadores internacionais para que prestem mais assistência face à destruição e danos causados pelo ciclone tropical Freddy" (Lusa2_182).

(160) (...) o trânsito esteve cortado nos dois sentidos, ao quilómetro 272 da EN 18, desde cerca das 17:30, após uma colisão rodoviária que envolveu três veículos ligeiros de passageiros e um ligeiro de mercadorias (Lusa2_158).

No exemplo (157), infere-se uma relação causal implícita a partir do léxico *tiroteio*, correspondente ao Arg.2 (Razão), que sugere uma ação suscetível de originar a consequência expressa no Arg.1 (Resultado) (*morreram 13 pessoas*). No exemplo (158), a relação causal implícita é inferida a partir do conteúdo proposicional (Cunha & Silvano, 2009). O segmento *quando ocorreu o desabamento*, correspondente ao Arg.2 (Razão), funciona como explicação para *o morador (...) sofreu escoriações*, Arg.1 (Resultado). Apesar da ausência de um marcador explícito de causalidade, a relação é inferida a partir do conteúdo das orações, ligadas pela conjunção *quando*, que por ter um valor neutro pode ter associada uma relação intrínseca, como é o caso da relação causal (Cunha & Silvano, 2009). No exemplo (159), as frases apresentam independência sintática, mas é possível inferir uma relação causal implícita entre elas. A primeira frase, correspondente ao Arg.2 (Razão) *uma tragédia nacional*, constitui o motivo do Arg.1 (Resultado) *Apelo aos parceiros e doadores internacionais para que prestem mais assistência face à destruição e danos causados pelo ciclone tropical Freddy*. Embora não haja um marcador explícito de causalidade, o apelo à assistência surge como consequência da tragédia

nacional. No exemplo (160), não se observa um marcador explícito de causalidade. Contudo, a relação causal implícita é inferida a partir da semântica composicional, incluindo da informação temporal. O Arg.2 (Razão) *após uma colisão rodoviária que envolveu três veículos ligeiros de passageiros e um ligeiro de mercadorias* explica o Arg.1 (Resultado) *o trânsito esteve cortado nos dois sentidos, ao quilómetro 272 da EN 18, desde cerca das 17:30.*

Para além disso, as relações causais podem ocorrer em orações adverbiais participiais, gerundivas e infinitivas, simples ou compostas, como se observa nos exemplos (161), (162) e (152).

(161) duas pessoas ficaram feridas, **atingidas** a golpes de martelo (Lusa2_39).

(162) Muitas pessoas saíram para as ruas dessas cidades, **temendo** que as casas desabassem (Lusa2_219)

(163) "O Presidente da República lamenta as vítimas do terremoto desta manhã, **tendo enviado** uma mensagem solidária ao Presidente Erdogan da Turquia" (Lusa2_241)

Neste sentido, importa relembrar que as construções causativas, na língua inglesa, se apresentam sob duas formas, segundo Girju (2003), relações explícitas e relações implícitas. De acordo com o autor,

explicit causation patterns can contain relevant keywords such as *cause, effect, consequence*, but also ambiguous ones such as *generate, induce*, etc. The implicit causative constructions are more complex, involving inference based on semantic analysis and background knowledge. (Girju, 2003, p. 3).

Os resultados obtidos revelam a predominância de relações causais explícitas nas notícias em português europeu, o que se articula com a natureza essencialmente informativa e objetiva do discurso jornalístico. Com efeito, ao procurar transmitir a informação de forma clara e acessível, recorrendo a um registo de língua corrente e evitando ambiguidades interpretativas, o jornalista privilegia a explicitação das relações causais por meio de marcadores linguísticos que orientam o leitor na identificação dessas relações. Este aspeto é particularmente relevante se se considerar que a compreensão e a interpretação do discurso não dependem exclusivamente da

competência linguística do falante, mas também do conhecimento do mundo e da capacidade de mobilizar informação contextual e gramatical específica da língua (Van den Broek, 1989). Quanto às relações causais explícitas, observa-se ainda a mobilização de diferentes categorias gramaticais que contribuem para a codificação da causalidade. Tal como sublinham Khoo et al. (2002), existe uma grande diversidade de expressões linguísticas capazes de sinalizar explicitamente relações de causa e efeito. Na língua portuguesa, este fenómeno manifesta-se de forma particularmente evidente através do uso de marcadores ou articuladores discursivos, que podem assumir a forma de advérbios, partículas, conjunções coordenativas e subordinativas, bem como resultar de configurações específicas da ordem das palavras (Couper-Kuhlen & Kortmann, 2000).

6.2.2. Classificação dos sinais de causa

Na sequência da análise das relações explícitas, focamos a nossa atenção nos marcadores dessas relações, isto é, os elementos linguísticos explícitos que marcam a relação de causalidade, introduzindo o argumento de razão ou de resultado. De acordo com a Norma ISO 24617-8 (2016), estes elementos, podem ser conetores, que constituem elementos fundamentais na estruturação e organização textual, uma vez que asseguram a articulação e a progressão do discurso (Lopes & Carrilho, 2020), permitindo uma interpretação clara e coerente da informação veiculada no texto, mas também outros nomes, verbos e preposições com valor causativo.

Como observado anteriormente, as relações de causalidade nas notícias em português europeu tendem a ser objetivas e expressas de forma explícita através de elementos linguísticos específicos. Estes elementos têm sido amplamente estudados por diversos autores, como Mirza et al. (2014), que analisaram classes como preposições, conjunções, conetores adverbiais e expressões integradas na oração; Prasad et al. (2010), que se debruçaram sobre sinais de causa como conjunções coordenativas, sintagmas preposicionais e advérbios; Girju (2003), que classificou as marcas de causalidade em causativos simples, resultativos e instrumentais; Lopes (2004), que desenvolveu um estudo sobre conetores discursivos causais e explicativos; e Silvano

(2010), que analisou a conjunção *porque* em diferentes contextos de introdução da relação causal.

Com base nestes estudos, o presente trabalho classifica os elementos linguísticos que indicam uma relação de causa nas seguintes categorias: verbo causativo; preposição (simples e locuções prepositivas); nome; conjunção (simples e locuções conjuntivas); adjetivo; e advérbio. Apresenta-se, de seguida, a frequência dos sinais de causa na seguinte Tabela.

Tabela 17 - Frequência e classificação dos sinais de causa

SIGNAL CAUSAL	CATEGORIA	OCORRÊNCIA	PERCENTAGEM
	<i>Noun</i>	3	0,45%
	<i>Verb</i>	220	33,28%
	<i>Conjunction</i>	42	6,35%
	<i>Adverb</i>	1	0,17%
	<i>Adjective</i>	7	1,05%
	<i>Preposition</i>	388	58,70%
TOTAL	661		100%

Esta tabela oferece uma visão global da frequência dos sinais causais presentes nas relações causais explícitas nas notícias em português europeu. Observa-se que a classe³ mais frequente é a das preposições, com 388 ocorrências (58,70%), seguida dos verbos, com 220 ocorrências (33,28%). As conjunções registam 42 ocorrências (6,35%), os adjetivos surgem em 7 casos (1,05%), os nomes em 3 casos (0,45%) e, por fim, os advérbios apresentam uma expressão quantitativa muito reduzida, com apenas 1 ocorrência (0,17%).

Estes dados evidenciam uma forte tendência para que as relações causais nas notícias em português europeu sejam introduzidas sobretudo por preposições (Dirven, 1995), verbos (Girju, 2003) e conjunções (Lobo, 2003). As categorias de nome (Brito & Oliveira, 1997), adjetivo e advérbio (Lopes & Carrilho, 2020) revelam uma expressão quantitativa bastante reduzida na marcação da causalidade. Tal resultado confirma a

³ A classificação dos elementos linguísticos analisados segue fontes como dicionários (ex. Dicionário da Academia de Ciências de Lisboa), gramáticas de referência (ex. Gramática do Português, da Fundação Calouste Gulbenkian) e trabalhos dos diversos autores referidos ao longo deste trabalho.

observação de Asher & Lascarides (2003), segundo a qual o léxico causal que estrutura explicitamente a interpretação e a inferência das relações discursivas de causa se manifesta principalmente através de verbos causativos, nominalizações causativas e preposições.

Relativamente à classe das preposições, importa salientar que esta não se limita às preposições simples, abrangendo igualmente as locuções prepositivas, que desempenham a mesma função discursiva (Yu, 2023). Os exemplos seguintes ilustram estas ocorrências.

- (164) Um homem ficou em prisão preventiva **por** estar fortemente indiciado da prática de "um número significativo" de furtos em estabelecimentos comerciais e veículos no centro da cidade de Ponta Delgada, nos Açores (Lusa2_281)
- (165) o incêndio alastrou **por causa do** vento (Lusa2_196)
- (166) A partir de então, passou a ser vigiado pelos serviços secretos franceses **devido ao** seu perfil ligado ao radicalismo islâmico. (Lusa2_39)
- (167) O Governo pediu para reforçar a vigilância "**face ao** risco de deslizamentos de terra, inundações e rios cheios", na sequência da tempestade tropical Mawar, anteriormente classificada como tufão. (Lusa2_199)
- (168) Um dos emblemas da Serra Leoa, a secular "Árvore do Algodão" da capital, Freetown, perdeu todos os ramos **na sequência de** uma tempestade na quarta-feira à noite (Lusa2_198)
- (169) As autoridades japonesas apelaram à prudência durante a continuação do fenómeno, **dado** o risco de tempestades, ventos fortes, ondas e aluimentos de terra. (Lusa2_206)
- (170) Portugal enviou hoje um grupo de 100 operacionais e cerca de 30 viaturas **em resposta a** um pedido de assistência, para combate a incêndios florestais, feito pela Proteção Civil espanhola. (Lusa2_196)
- (171) os dois homens morreram hoje num acidente de trabalho na fábrica da Repsol Polímeros, em Sines, **ao** caírem de "uma altura de mais de 30 metros" (Lusa2_350)

(172) "Fiquei chocada **ao** ver o nosso querido algodoeiro destruído quando ia para o trabalho esta manhã"(Lusa2_198)

No que diz respeito à classe dos verbos, destacam-se as formas verbais conjugadas, ilustradas nos exemplos (173) a (177), bem como as formas não finitas, nomeadamente o infinitivo (178), o particípio passado (179) e o gerúndio (180). Incluem-se ainda construções na voz passiva dos verbos causativos (181) e construções perifrásticas, resultantes da combinação do verbo auxiliar *estar* com a preposição simples *a* e o verbo causativo no infinitivo, como em (182).

(173) A forte neve **obrigou** também ao adiamento do jogo da Bundesliga entre o Bayern de Munique e o Union Berlin. (Lusa2_29)

(174) a investigação policial **permitiu** apurar que se tratava de uma situação relacionada com tráfico de estupefacientes (Lusa2_356)

(175) O calor que atingiu grande parte dos Estados Unidos esta sexta-feira **forçou** ao encerramento de algumas escolas em Nova Inglaterra (Lusa2_218)

(176) Em Madrid, as chuvas torrenciais **levaram** também ao encerramento temporário de várias linhas de metro (Lusa2_223)

(177) Chuvas torrenciais, ventos até 100 quilómetros por hora, inundações e, em alguns casos, tempestades de granizo, **deixaram** mais de 124 mil pessoas sem eletricidade no sul de Queensland no dia de Natal. (Lusa2_119)

(178) depois de as chuvas torrenciais terem **provocado** grandes deslizamentos de terra em Bitanga, Bigombe e Mazozo"(Lusa2_117)

(179) as dezenas de mortes **provocadas** por chuvas torrenciais No norte daquele país (Lusa2_111)

(180) A forte precipitação na segunda-feira (...) **afetando** ainda a circulação rodoviária na cidade e **obrigando** ao encerramento de dezenas de estações de metro (Lusa2_109)

(181) estes deslizamentos de terras foram também **provocados** pela cheia de um rio (Lusa2_117)

(182) Uma derrocada ocorrida esta manhã está a **causar** constrangimentos na Estrada Nacional 2 (EN2), junto ao Peso da Régua. (Lusa2_132).

No que diz respeito à classe das conjunções, que representa 6,35% das ocorrências, identificaram-se relações causais introduzidas por subordinativas, como em (184), e por locuções conjuncionais, como em (185).

(183) o número de mortos pode aumentar **porque** as equipas de socorro continuam a remover as toneladas de terra que caíram sobre os condutores e passageiros.

(184) os suspeitos se ausentavam “de forma suspeita, **pelo que** vieram a ser intercetados” (Lusa2_307)

Contrariamente ao esperado, os resultados desta investigação mostram que, em português europeu nas notícias da Agência Lusa, os sinais de causa mais frequentes pertencem às classes das preposições e dos verbos, e não às conjunções ou advérbios conetivos. Este trabalho mostra como as preposições e locuções prepositivas podem funcionar como pistas relevantes para a inferência de relações discursivas, em particular das relações causais (Dirven, 1995). Conforme referido no quadro concetual apresentado no capítulo 3, as preposições exprimem relações de proximidade, origem e envolvimento ou contenção, sendo estes esquemas espaciais a base para a derivação da interpretação causal. A análise realizada revela ainda o recurso frequente a um conjunto de verbos que, segundo Soares da Silva (1999), incorporam a noção de causa ou causalidade na sua significação, ao expressarem a produção de uma ação, a aquisição de uma propriedade ou a entrada num determinado estado. O uso destes elementos linguísticos para assinalarem de forma explícita a causalidade manifesta-se de forma clara nas notícias em português europeu analisadas neste corpus, permitindo o estabelecimento de relações de causa de modo objetivo e informativamente eficaz.

6.2.3. Propriedades dos eventos nas relações de causa

Este trabalho de investigação identificou também propriedades que caracterizam os eventos que compõem uma relação de causa. Das 829 relações causais registadas nas 300 notícias, contabilizaram-se 1804 eventos nos argumentos de razão e de resultado. Neste sentido, procurou-se identificar: (i) os atributos e valores dos eventos nas 829 relações causais (Tabela 18), (ii) os atributos e valores dos eventos no argumento de razão e no argumento de resultado (Tabela 19) e (iii) as combinações dos valores dos eventos associados aos argumentos de razão e de resultado (Tabela 20).

Pretendeu-se, assim, analisar o que acontece ou ocorre numa relação causal, eventos ou estados, recorrendo a atributos e respetivos valores, tal como definidos no esquema de anotação que propusemos (LUSA-ACT), como *Class* (occurrence, state, reporting, perception, *I_Action*, *I_State*, aspectual), *Event_Type* (state, process, transition), *Pos* (verb, noun, adjective, preposition), *Tense* (past, present, future), *Aspect* (progressive, perfective, imperfective), *Vform* (gerundive, infinitive, participle), *Mood* (subjunctive, conditional, future), *Modality* (dever, poder, ter de) e *Polarity* (negative, positive).

Relativamente aos atributos e valores dos 1804 eventos identificados nas relações causais das notícias, apresentam-se, na tabela seguinte, as respetivas ocorrências e percentagens.

Tabela 18 - Atributos e valores dos eventos nas relações causais

EVENT			
ATRIBUTO	VALOR	OCORRÊNCIA	PERCENTAGEM
Class	<i>Occurrence</i>	1421	78.77%
	<i>State</i>	280	15.52%
	<i>Reporting</i>	10	0.55%
	<i>I_State</i>	50	2.77%
	<i>I_Action</i>	41	2.27%
	<i>Perception</i>	2	0.12%
	TOTAL	1804	100%
Event_Type	<i>Transition</i>	641	74.19%
	<i>State</i>	205	23.73%
	<i>Process</i>	18	2.08%
	TOTAL	864	100%
Pos	<i>Verb</i>	715	39.03%
	<i>Noun</i>	960	53.33%

	<i>Adjective</i>	102	5.65%
	<i>Preposition</i>	27	1.50%
TOTAL		1804	100
Tense	<i>Past</i>	489	75.35%
	<i>Present</i>	148	22.80%
	<i>Future</i>	12	1.85%
	TOTAL	649	100%
Aspect	<i>Perfective</i>	548	92.41%
	<i>Imperfective</i>	27	4.55%
	<i>Progressive</i>	17	2.87%
	<i>Imperfective_Progressive</i>	1	0.17%
	TOTAL	593	100%
Vform	<i>Participle</i>	13	6.31%
	<i>Infinitive</i>	140	67.96%
	<i>Gerundive</i>	53	25.73%
	TOTAL	206	100%
Mood	<i>Future</i>	11	66.11%
	<i>Conditional</i>	4	22.22%
	<i>Subjunctive</i>	3	16.67%
	TOTAL	18	100%
Modality	<i>Poder</i>	18	62.07%
	<i>Dever</i>	1	3.45%
	<i>Ter_de</i>	10	34.48%
	TOTAL	29	100%
Polarity	<i>Neg</i>	30	1.66%
	<i>Pos</i>	1774	98.34%
	TOTAL	1804	100%

Conforme se observa na tabela, as propriedades dos eventos identificados no corpus apresentam alguma variedade. Considerando o valor mais frequente em cada atributo, verifica-se que, em *Class*, predomina *Occurrence* (78,77%); em *Event_Type*, destaca-se *Transition* (74,19%); em *Pos*, prevalece *Noun* (53,33%); em *Tense*, predomina *Past* (75,35%); em *Aspect*, destaca-se *Perfective* (92,41%); em *Vform*, é mais frequente *Infinitive* (67,96%); em *Mood*, predomina *Future* (66,11%); em *Modality*, destaca-se *poder* (62,07%); e, em *Polarity*, predomina *positive* (98,34%).

Estes resultados indicam que a maioria das situações nas relações causais é identificada como evento, uma situação dinâmica, durativa e télica. Regista-se ainda uma percentagem de eventualidades categorizadas como estados, de natureza não dinâmica (Cunha, 2013). Observem-se os seguintes exemplos ilustrativos.

(185) O **terramoto** de magnitude 5,9 causou **danos** generalizados nas províncias de Gansu e Qinghai, incluindo o desabamento de casas, forçando os moradores a **fugir** para as ruas (Lusa2_114)

(186) Este tipo de incidente é **comum** aqui em Masisi devido à **fragilidade** do solo.(Lusa2_185)

No exemplo (185), identificam-se duas relações causais. Na primeira, o Arg.2 (Razão) é *O terremoto de magnitude 5,9* e o Arg.1 (Resultado) é *danos generalizados nas províncias de Gansu e Qinghai, incluindo o desabamento de casas*. Além disso, o mesmo Arg.2 (Razão) estabelece uma segunda relação causal com o Arg.1 (Resultado) *os moradores a fugir para as ruas*. A partir destas duas relações, identificam-se quatro eventos principais: terramoto (contabilizado duas vezes), com valores (Occurrence, Noun, Pos); danos (Occurrence, Noun, Pos); e fugir (Occurrence, Transition, Verb, Infinitive, Pos). Estes acontecimentos apresentam natureza durativa e dinâmica. No exemplo (186), identificou-se uma relação causal entre o Arg.2 (Razão) *à fragilidade do solo* e o Arg.1 (Resultado) *Este tipo de incidente é comum aqui em Masisi*. Com base nessa relação causal, ambos eventos principais são classificados como estados: *comum* (State, State, Adjective, Present, Pos) e *fragilidade* (State, State, Noun, Pos).

Quanto aos atributos e respectivos valores nos argumentos de razão e de resultado, apresenta-se, de seguida, uma breve comparação, com base na seguinte Tabela.

Tabela 19 - Atributos e valores dos eventos no argumento razão e resultado

EVENTOS NO ARGUMENTO RAZÃO E NO RESULTADO			
ATRIBUTO	VALOR	RAZÃO	RESULTADO
Class	<i>Occurrence</i>	735/80,42%	686/77,08%
	<i>State</i>	155/16,96%	125/14,04%
	<i>Reporting</i>	1/0,11%	9/1,01%
	<i>I_State</i>	15/1,64%	35/3,93%
	<i>I_Action</i>	7/0,77%	34/3,82%
	<i>Perception</i>	1/0,11%	1/0,11%
	Total	914/100%	100/100%

Event_Type	<i>Transition</i> <i>State</i> <i>Process</i>	159/70,40% 63/27,75% 5/2,20%	482/75,66% 142/22,29% 13/2,04%
	Total	227/100%	637/100%
Pos	<i>Verb</i> <i>Noun</i> <i>Adjective</i> <i>Preposition</i>	188/20,57% 696/76,15% 24/2,63% 6/0,66%	527/59,21% 264/29,66% 78/8,76% 21/2,36%
	Total	914/100%	890/100%
Tense	<i>Past</i> <i>Present</i> <i>Future</i>	77/65,25% 40/33,90% 1/0,85%	412/77,59% 108/20,34% 11/2,07%
	Total	118/100%	531/100%
Aspect	<i>Perfective</i> <i>Imperfective</i> <i>Progressive</i> <i>Imperfective_Progressive</i>	122/92,42% 8/6,06% 2/1,52%	426/92,41% 19/4,12% 15/3,25% 1/0,22%
	Total	138/100%	461/100%
Vform	<i>Participle</i> <i>Infinitive</i> <i>Gerundive</i>	6/5,88% 86/84,31% 10/9,80%	7/6,73% 54/51,92% 43/41,35%
	total	102/100%	104/100%
Mood	<i>Future</i> <i>Conditional</i> <i>Subjunctive</i>	1/33,33% 2/66,67%	10/66,67% 2/13,33% 3/20,00%
	Total	3/100%	15/100%
Molarity	<i>Poder</i> <i>Ter_de</i> <i>Dever</i>	9/100%	9/45,00% 10/50,00% 1/5,00%
	Total	9/100%	20/100%
Polarity	<i>Neg</i> <i>Pos</i>	15/1,64% 899/98,36%	15/1,69% 875/98,31%
	Total	914/100%	890/100%

A tabela mostra que os valores dos eventos nos argumentos apresentam semelhanças em alguns atributos e diferenças noutros. Por exemplo, em *Class*, *Occurrence* predomina tanto no argumento de razão (80,42%) como no de resultado (77,08%). Observa-se igualmente predominância semelhante em *Event_Type* (*Transition*: 70,40% e 75,66%), em *Tense* (*Past*: 65,25% e 77,59%), em *Aspect* (*Perfective*: 92,42% e 92,41%), em *Vform* (*Infinitive*: 84,31% e 51,92%) e em *Polarity* (*Pos*: 98,36% e 98,31%). Quanto às diferenças, destacam-se os atributos *Pos*, em que o argumento de razão apresenta predominância

de *Noun* (76,15%) e o de resultado de *Verb* (59,21%); *Mood*, em que se registam *Conditional* no argumento de razão (2 ocorrências/66,67%) e *Future* no argumento de resultado (10 ocorrências/66,67%); e *Modality*, com *poder* no argumento de razão (9 ocorrências/100%) e *ter_de* no argumento de resultado (10 ocorrências/ 50,00%).

Após esta comparação, procede-se à análise das combinações (pares) dos valores dos eventos associados aos argumentos de razão e de resultado em cada relação causal. Importa referir que se optou por destacar as combinações com frequência superior a 5%, agregando as restantes na categoria *Outros*.

Tabela 20 - Combinações dos valores dos eventos dos argumentos de razão e resultado

ATRIBUTOS	COMBINAÇÕES DOS VALORES	OCORRÊNCIA
CLASS	(Reason, Occurrence), (Result, Occurrence)	625 / 65.24%
	(Reason, Occurrence), (Result, State)	109 / 11.38%
	(Reason, State), (Result, Occurrence)	106 / 11.07%
	Outros	118 / 12.32%
EVENT_TYPE	(Reason, None), (Result, Transition)	352 / 35.92%
	(Reason, None), (Result, None))	273 / 27.86%
	(Reason, None), (Result, State)	113 / 11.53%
	(Reason, Transition), (Result, Transition)	123 / 12.55%
	Outros	119 / 12.14%
POS	(Reason, Noun), (Result, Verb)	390 / 39.80%
	(Reason, Noun), (Result, Noun)	283 / 28.88%
	(Reason, Verb), (Result, Verb)	160 / 16.33%
	(Reason, Noun), (Result, Adjective)	66 / 6.73%
	Outro	81 / 8.27%
TENSE	(Reason, None), (Result, Past)	392 / 40.00%
	(Reason, None), (Result, None)	354 / 36.12%
	(Reason, None), (Result, Present)	103 / 10.51%
	Outros	131 / 13.37%
ASPECT	(Reason, None), (Result, Perfective)	445 / 45.41%
	(Reason, None), (Result, None))	372 / 37.96%
	(Reason, Perfective), (Result, Perfective)	85 / 8.67%
	Outros	78 / 7.96%
VFORM	(Reason, None), (Result, None))	779 / 79.49%
	(Reason, Infinitive), (Result, None)	77 / 7.86%
	(Reason, None), (Result, Infinitive)	52 / 5.31%
	Outros	72 / 7.35%
MOOD	(Reason, None), (Result, None)	959 / 97.86%
	Outros	21 / 2.14%

MODALITY	(Reason, None), (Result, None)	949 / 96.84%
	Outros	31 / 3.16%
POLARITY	(Reason, Pos), (Result, Pos)	944 / 96.33%
	Outro	36 / 3.67%

As combinações mais frequentes registadas no corpus evidenciam padrões claros. Em *Class*, a combinação dominante é (Reason, Occurrence) – (Result, Occurrence), com 625 ocorrências (65,24%), indicando que os eventos de razão e de resultado são, maioritariamente, dinâmicos em ambos os argumentos. Em *Event_Type*, a combinação mais frequente é (Reason, None) – (Result, Transition), com 352 ocorrências (35,92%), sugerindo uma forte presença de eventos expressos por nomes no argumento de razão, sem caracterização de *Event_Type*, e eventos dinâmicos, durativos e télicos no argumento de resultado. Em *Pos*, a combinação mais utilizada é (Reason, Noun) – (Result, Verb), com 390 ocorrências (39,80%), evidenciando que o evento da razão tende a ser representado por nomes e o do resultado por verbos. Em *Tense*, destaca-se a combinação (Reason, None) – (Result, Past), com 392 ocorrências (40,00%), mostrando novamente que os eventos de razão são frequentemente nomes, e, por isso, não apresentam informação temporal, enquanto os eventos de resultado surgem maioritariamente em verbos conjugados no passado. Em *Aspect*, prevalece (Reason, None) – (Result, Perfective), com 445 ocorrências (45,41%), revelando a caracterização perfeitiva dos eventos de resultado. Quanto a *Vform*, observa-se uma elevada frequência de (Reason, None) – (Result, None), com 779 ocorrências (79,49%), o que indica uma predominância de eventos representados por formas finitas. Nos atributos *Mood* e *Modality*, regista-se uma distribuição quase absoluta do par (Reason, None) – (Result, None), com 959 ocorrências (97,86%) e 949 ocorrências (96,84%), respetivamente, o que sugere um uso pouco frequente de modo e modalidade nos eventos associados às relações causais, em consonância com a caracterização informativa e factual da notícia. Por fim, em *Polarity*, a combinação (Reason, Pos) – (Result, Pos) é dominante, com 944 ocorrências (96,33%), evidenciando a predominância de eventos afirmativos. O exemplo seguinte ilustra as propriedades mais frequentes dos eventos principais representados pelos argumentos de razão e de resultado.

(187) Em agosto de 2017, mais de 1.100 pessoas morreram na sequência de um deslizamento de lama na capital causado por fortes chuvas. (Lusa2_198)

No exemplo (187), o evento no argumento de resultado é *morreram*, caracterizado como *Occurrence, Transition, Verb, Past, Perfective* e *Pos*. O evento no argumento de razão é *deslizamento*, caracterizado como *Occurrence, Noun* e *Pos*. Neste caso, as combinações dos valores entre os dois eventos (Resultado, Razão) seriam: Class (*Occurrence, Occurrence*), Event_Type (*Transition, Noun*), Pos (*Verb, Noun*), Tense (*Past, Noun*), Aspect (*perfective, Noun*) e Polarity (*Pos, Pos*).

De um modo geral, os resultados relativos às propriedades dos eventos identificados podem ser diretamente articulados com as características do discurso noticioso. A configuração mais frequente dos eventos principais que integram os argumentos de razão e de resultado corresponde a situações de natureza dinâmica, localizadas no passado e perspectivadas como terminadas, o que é consistente com o facto de as notícias relatarem predominantemente acontecimentos já ocorridos. Para além disso, a presença recorrente de eventos do tipo *Transition*, nomeadamente processos culminados e culminações, no passado favorece leituras de sucessividade temporal, típicas de estruturas narrativas. Esta característica revela-se particularmente relevante num género discursivo como a notícia, que frequentemente organiza a informação segundo uma progressão temporal dos acontecimentos. Ainda assim, a ocorrência de eventos classificados como estados evidencia que as relações causais em notícias não se limitam a situações dinâmicas. Para além disso, estas características vão ao encontro da natureza narrativa das notícias que constituem o nosso corpus. Tal como refere da Silva (2005), o predomínio de situações eventivas representadas no Pretérito Perfeito caracterizam sequências predominantemente narrativas.

6.2.4. Relações temporais entre dois argumentos das relações causais

Sendo as propriedades temporais relevantes para a caracterização semântica das relações causais, no nosso esquema de anotação, LUSA-ACT, incluímos etiquetas, atributos e valores que nos permitissem representar a informação temporal. Na secção

anterior apresentamos a análise das características dos eventos e nesta secção focamos a nossa atenção nas relações temporais que se estabelecem entre os eventos principais dos argumentos de razão e resultado.

Como já foi referido na secção dedicada ao processo de anotação, as relações temporais estabelecem-se entre os dois eventos que integram os argumentos das relações de causa, com direcionalidade do posterior para o anterior na ordem linear do discurso. Foram considerados os tipos *Before*, *After*, *Simultaneous*, *Includes* e *Is_includes*, bem como as respetivas ocorrências no corpus. A seguinte Tabela sistematiza a frequência dos tipos de relações temporais identificados nas relações causais.

Tabela 21 - Tipos de relações temporais na relações causais

RELAÇÕES TEMPORAIS ENTRE OS EVENTOS DOS DOIS ARGUMENTOS DAS RELAÇÕES CAUSAIS	ATRIBUTOS	OCORRÊNCIA	PERCENTAGEM
	<i>Before</i>	440	53,08%
	<i>After</i>	250	30,16%
	<i>Simultaneous</i>	41	4,95%
	<i>Includes</i>	69	8,32%
	<i>Is_Included</i>	29	3,49%
TOTAL	829		100%

Da análise das 829 relações causais, observa-se predominância da relação temporal *Before*, estabelecida entre os dois eventos que integram os argumentos de razão e resultado, com 440 ocorrências (53,08%). Segue-se a relação *After*, com 250 casos (30,16%). Quanto à relação *Simultaneous*, registam-se 41 casos (4,46%). No tipo *Includes*, verificam-se 69 ocorrências (8,32%) e no tipo *Is_included*, registam-se 29 casos (3,49%). Para ilustrar estas ocorrências, apresentam-se os seguintes exemplos de cada tipo de relação temporal.

(188) Três ativistas do movimento Climáximo **foram** hoje detidos pela PSP após terem **bloqueado** o trânsito durante cerca de meia hora a Rua da Escola Politécnica, em Lisboa. (Lusa2_144)

(189) A **queda** do carvalho provocou **13** mortes (duas das quais de cidadãos estrangeiros, de nacionalidades francesa e húngara) (Lusa2_362)

- (190) As piscinas de Águas Santas, na Maia, estão **encerradas** há uma semana por **suspeita** da existência da bactéria da ‘legionella’ (Lusa2_137)
- (191) No âmbito da investigação, foi ainda **constituída arguida** a mulher do detido, por **suspeitas** de coautoria em alguns dos crimes. (Lusa2_363)
- (192) No período em que o Teatro Camões, casa da CNB, se encontra **encerrado** para obras de reestruturação no âmbito do Plano de Recuperação e Resiliência (PRR), a companhia nacional **fez** uma digressão para apresentar este clássico do romantismo, começando pelo Coliseu do Porto, depois, Coimbra, no Convento São Francisco, em ambos os espaços com a atuação da Orquestra Clássica do Centro, e seguidamente no Teatro das Figuras, em Faro, com a Orquestra do Algarve

No exemplo (188), identifica-se uma relação de anterioridade (*Before*): o Arg.1 (Resultado) *Três ativistas do movimento Climáximo foram hoje detidos pela PSP* ocorre posteriormente ao Arg.2 (Razão) *após terem bloqueado o trânsito durante cerca de meia hora a Rua da Escola Politécnica, em Lisboa*. No exemplo (189), identifica-se uma relação de *After* entre o Arg.2 (Razão) *A queda do carvalho* e o Arg.1 (Resultado) *13 mortes (duas das quais de cidadãos estrangeiros, de nacionalidades francesa e húngara)*, assinalada pelos eventos principais *queda* e *mortes*. Observa-se que o evento *mortes* ocorre após o evento *queda*, o que configura uma relação temporal de posterioridade. No exemplo (190), pode inferir-se que a relação temporal entre os dois eventos é de *Simultaneous*. Neste caso, ambos eventos correspondem a estados, o evento *encerradas* ocorre ao mesmo tempo que o evento *suspeita*. No exemplo (191), o intervalo temporal associado ao estado de ser suspeito inclui (*Includes*) o acontecimento expresso por *constituída arguida* no argumento de resultado. Por último, no exemplo (192) é ilustrada a relação de *Is_Included*: o acontecimento do evento *fez* está incluído no tempo de intervalo do evento *encerrada*.

Portanto, as relações temporais mais frequentes entre os argumentos das relações causais identificadas em notícias de natureza narrativa podem ser *Before*, *After* e *Simultaneous*, como observado por Silvano (2010) nas frases com orações adverbiais causais.

Analisaram-se também as relações temporais em relações causais explícitas e implícitas. No conjunto de dados selecionado, foram identificados vários tipos nas relações causais explícitas, conforme apresentado na seguinte Tabela.

Tabela 22 - Relações temporais nas relações causais explícitas

RELAÇÕES TEMPORAIS	RELAÇÕES CAUSAIS EXPLÍCITAS	RELAÇÕES CAUSAIS EXPLÍCITAS
<i>Before</i>	366/55,37%	74/44,05%
<i>After</i>	198/29,95%	52/30,95%
<i>Simultaneous</i>	29/4,39%	19/11,31%
<i>Includes</i>	50/7,56%	12/7,14%
<i>Is-included</i>	18/2,73%	11/6,55%
TOTAL	661/100%	168/100%

A partir dessa tabela, verifica-se que, quer nas relações causais explícitas, quer nas implícitas, a relação temporal *Before* ocorre com maior frequência, com 366 casos (55,37%) e com 74 ocorrências (44,05%), respectivamente. Segue-se *After*, com 198 ocorrências (29,95%), no caso das relações explícitas, e 52 casos (30,95%), no caso das implícitas. A terceira mais frequente nas relações explícitas é *Includes*, com 50 ocorrências (7,56%), enquanto nas relações implícitas é a de *Simultaneous*, com 19 ocorrências (11,31%). Os exemplos seguintes ilustram dois destes casos.

(193) As **altas temperaturas** no início da semana causaram **interrupções** nas escolas desde Michigan até à Virgínia, com alguns distritos a dispensarem alunos mais cedo (Lusa2_218)

(194) A estrada que liga Medellín, a segunda maior cidade da Colômbia, a Quibdó, na região de Choco, foi temporariamente **fechada** depois de ter sido **danificada** por vários deslizamentos de terra (Lusa2_126)

No primeiro exemplo, identifica-se uma relação causal marcada por *causaram*, em que a relação temporal entre os eventos *altas temperaturas* e *interrupções* é *After*. No segundo exemplo, identifica-se uma relação causal implícita, em que a relação temporal entre os eventos *fechada* e *danificada* é *Before*.

Como referido anteriormente, a ordem dos argumentos nas relações causais tende a seguir as configurações Arg.2 (Razão) – Arg.1 (Resultado) ou Arg.1 (Resultado) – Arg.2 (Razão). Esta ordenação pode influenciar as relações temporais, pelo que esta análise também foi feita.

Tabela 23 - Ordem dos argumentos nas relações temporais

RELAÇÕES TEMPORAIS	ARG.2 (RAZÃO)- ARG.1(RESULTADO)	ARG.1(RESULTADO) - ARG.2 (RAZÃO))
<i>Before</i>	14/4,62%	426/80,99%
<i>After</i>	242/79,87%	8/1,52%
<i>Simultaneous</i>	18/5,94%	23/4,37%
<i>Includes</i>	2/0,66%	67/12,74%
<i>Is-included</i>	27/8,91%	2/0,38%
TOTAL	303/100%	526/100%

Nos casos em que, na ordem linear do discurso, o Arg.2 (Razão) ocorre antes do Arg.1 (Resultado), verifica-se predominância clara da relação *After*, com 242 ocorrências (79,87%), um valor substancialmente superior ao de *Before*, com 14 casos (4,62%). *Simultaneous*, *Is_included* e *Includes* apresentam frequências reduzidas, com 18 (5,94%), 27 ocorrências (8,91%) e 2 casos (0,66%), respetivamente. Quando a ordem é Arg.1 (Resultado) – Arg.2 (Razão), os dados revelam que *Before* é a relação temporal predominante, com 426 ocorrências (80,99%). *After* surge apenas 8 vezes (1,52%) e *Simultaneous* contabiliza 23 casos (4,37%). *Includes* e *Is_included* registam 67 (12,74%) e 2 casos (0,38%), respetivamente. Os exemplos seguintes ilustram a tendência.

(195) A **chegada** do ciclone obrigou ao **encerramento** de mais de uma centena de escolas e ao cancelamento de voos no estado australiano de Queensland, no norte do país. (Lusa2_130)

(196) Alguns **sofreram** queimaduras devido ao **calor intenso** (Lusa2_17)

No primeiro exemplo, a relação causal apresenta a ordem Razão–Resultado e a relação temporal entre os eventos *chegada* e *encerramento* é *After*. No segundo exemplo, a

relação causal apresenta a ordem Resultado–Razão e a relação temporal entre os eventos *queimaduras* e *calor* é *Before*.

Tabela 24 - Ordem dos argumentos das relações temporais nas relações causais explícitas

ORDEM DOS ARGUMENTOS	RELAÇÕES TEMPORAIS	RELAÇÕES CAUSAIS EXPLÍCITAS	RELAÇÕES CAUSAIS IMPLÍCITAS
ARG.2 (RAZÃO)- ARG.1(RESULTADO)	<i>Before</i>	10/4,42%	4/5,25%
	<i>After</i>	190/84,07%	51/67,11%
	<i>Simultaneous</i>	9/3,98%	9/11,84%
	<i>Includes</i>	1/0,45%	1/1,32%
	<i>Is-included</i>	16/7,08%	11/14,47%
TOTAL		226/100%	76/100%
ARG.1(RESULTADO) - ARG.2 (RAZÃO)	<i>Before</i>	357/82,07%	70/76,09%
	<i>After</i>	7/1,61%	1/1,08%
	<i>Simultaneous</i>	20/4,60%	3/3,26%
	<i>Includes</i>	49/11,26%	18/19,57%
	<i>Is-included</i>	2/0,46%	-
TOTAL		435/100%	92/100%

Os resultados da análise apresentados na Tabela anterior mostram que não há diferenças entre as relações causais implícitas e explícitas no que diz respeito à ordem dos argumentos e às relações temporais mais frequentes. Em ambos os casos, com a ordem Arg.2 (Razão) – Arg.1 (Resultado), a relação temporal mais frequente destaca-se *After*, e com a ordem Arg.1 (Resultado) – Arg.2 (Razão), verifica-se a predominância de *Before*.

Estes resultados confirmam que as relações temporais entre eventos em relações causais refletem uma característica de sucessividade na organização dos acontecimentos nas notícias.

Em trabalhos anteriores (Asher & Lascarides, 2003; Silvano, 2010, entre outros), a relação causal é frequentemente associada a uma relação temporal em que a razão precede o resultado. A nossa análise confirma esta tendência, que se coaduna também com a organização cronológica que as notícias devem seguir no relato dos eventos (Gonçalves, 2019; LUSA, 2019).

Identificadas as relações temporais mais frequentes nas diferentes configurações dos argumentos das relações causais, foi também nosso objetivo determinar que fatores linguísticos, no âmbito dos nossos parâmetros de análise, mais contribuem para a inferência das relações temporais. Para esta análise, consideramos apenas os eventos representados por verbos, por veicularem informação temporal e aspetual essencial para refletir sobre a influência de atributos como o tempo e a classe aspetual na determinação das relações temporais. Todos os pares combinatórios com menos de 5% foram agregados na categoria *Outros*. A Tabela seguinte sistematiza os resultados das associações que estabelecemos.

Tabela 25 - Combinações dos valores dos eventos associados a relações temporais

COMBINAÇÕES DOS VALORES DOS EVENTOS ASSOCIADOS AOS TIPOS DE RELAÇÕES TEMPORAIS					
ATRIBUTO	CLASS	EVENT_TYPE	TENSE	ASPECT	VFORME
BEF	(OCC,OCC)58/82,86% (ST, OCC)9/12,86% outros 3/4,29%	(TRAN, TRAN) 9/45% (ST, TRAN) 9/45% outros 2/10%	(PST, None) 41/58,57% (PST, PST) 10/14,29% (PRES, None) 6/8,57% (None, None) 5/ 7,14% outros 8/ 11,43%	(PERF, PERF) 35/50,0% (PERF, None)17/ 24,29% (None, PERF) 7/10,0% (None, None) 6/8,57% outros 5/ 7.14%	(None, INF) 44/62,86% (None, None) 17/24,29% outros 9/ 12,86%
AFT	(OCC, OCC)39/73,58% (OCC, IACT) 6/11,32% outros 8/15.09%	(TRAN, TRAN) 45/84,91% outros 8/15,09%	(PST, None) 17/32,08% (PST, PST) 12/22,64% (None, PST) 9/16,98% outros 15/28,30%	(PERF, PERF) 23/43,40% (PERF, None) 15/28,30% (None, PERF) 6/11,32% (None, None) 5/9,43% outros 4/7,55%	(None, None) 17/32,08% (None, GER) 15/28,30% (INF, None) 6/11,32% (None, INF) 5/9,43% outros 10/18,87%
INCL	(OCC, ST) 7/70,0% outros 3/30,0%	(TRAN, ST) 8/80,00% outros 2/20,00%	(PST, None) 4/40,0% (PRES, PRES) 3/30,0% (None, None) 2/20,0% (PST, PST) 1/10,0%	(PERF, None) 6/60,0% outros 3/40,0%	(None, None) 3/30,0% (None, INF) 3/30,0% (INF, None) 1/10,0% (GER, INF) 1/10,0% (None, GER) 1/10,0% (PART, INF) 1/10,0%
ISINCL	(ST, OCC) 3/50,0% (ST, IACT) 2/33,3% (OCC, OCC) 1/16,7%	(ST, TRAN) 5/83,33% outros 1/16,67%	(PRES, PST) 3/50,0% (PST, None) 1/16,67% (PRES, None) 1/16,67% (PST, PST) 1/16,67%	(None, PERF) 4/66,67% (PERF, PERF) 1/16,67% (IMPF, PERF) 1/16,67%	(None, None) 4/66,67% (None, GER) 2/33,33%
SIM	(ST, ST) 4/66,66% (OCC, OCC) 2/33,33%	(ST, ST) 2/66,66% (ST, PROC) 1/16,67% (TRAN, TRAN) 1/16,67%	(PST, None) 2/33,33% (PRES, PRES) 1/16,67% (PRES, None) 1/16,67% (PST, PST) 1/16,67% (None, PRES) 1/16,67%	(None, None) 3/50,0% (PERF, PERF) 1/16,67% (PERF, None) 1/16,67% (IMPF, None) 1/16,67%	(None, None) 2/33,33% (None, GER) 2/33,33% (None, INF) 1/16,67% (INF, None) 1/16,67%

Na relação temporal *BEFORE*, no atributo *Class*, a combinação mais frequente é (OCC, OCC), com 58 ocorrências (82,86%), indicando que a anterioridade se estabelece, maioritariamente, entre dois eventos dinâmicos. No atributo *Event_Type*, destacam-se os pares (TRAN, TRAN) e (ST, TRAN), que registam a mesma frequência (45 ocorrências; 84,91%), sugerindo que a relação temporal de anterioridade pode ocorrer entre dois eventos dinâmicos e télicos ou entre um estado e um evento télico. Quanto ao atributo *Tense*, predomina o par (PST, None), com 41 ocorrências (58,57%), o que indica que um dos eventos se localiza no passado, enquanto o outro não apresenta informação de localização temporal, neste caso, por se tratarem de formas não finitas (None). Relativamente ao atributo *Aspect*, a combinação dominante é (PERF, PERF), com 35 ocorrências (50,0%), mostrando que os dois eventos são apresentados como terminados, o que favorece relações de anterioridade, assim como de posterioridade. Por fim, no atributo *Vform*, a combinação mais frequente é (None, INF), com 44 ocorrências (62,86%), sugerindo que um dos eventos é frequentemente expresso por um verbo conjugado e o outro por um verbo no infinitivo.

Características semelhantes surgem na relação temporal *After*, diferenciando-se, contudo, em *Event_Type*, em que predomina (TRAN, TRAN), com 45 ocorrências (84,91%), e em *Vform*, em que a combinação mais frequente é (None, None), com 17 ocorrências (32,08%). No primeiro caso, ambos os eventos são caracterizados por situações que partilham propriedades aspetuais como o dinamismo e a telicidade; no segundo, observa-se que, na maioria dos casos, os eventos não são expressos por formas não finitas, mas por formas verbais conjugadas.

Nas relações *Includes* e *Is_included*, verifica-se um comportamento inverso nos atributos *Class* e *Event_Type*. De acordo com os resultados, na relação *Includes* registam-se os pares *Class* (OCC, ST) com 7 ocorrências (70,0%) e *Event_Type* (TRAN, ST) com 8 ocorrências (80,00%). Na relação *Is_included*, observam-se *Class* (ST, OCC) com 3 ocorrências (50,0%) e *Event_Type* (ST, TRAN) com 5 ocorrências (83,33%). Estas relações distinguem-se também em *Tense* (*Includes*: (PST, None) 4/40,0%; *Is_included*: (PRES, PST) 3/50,0%) e em *Aspect* (*Includes*: (PERF, None) 6/60,0%; *Is_included*: (None, PERF)

4/66,67%). Em contrapartida, apresentam semelhança no atributo *Vform*, com (None, None) na relação *Includes* (3/30,0%) e na relação *Is_included* (4/66,67%).

Na relação temporal *Simultaneous*, em que os dois eventos ocorrem ao mesmo tempo, registam-se as mesmas combinações nos atributos *Class* (ST, ST) (4/66,66%) e *Event_Type* (ST, ST) (4/66,66%), bem como em *Aspect* (None, None) (3/50,0%) e *Vform* (None, None) (2/33,33%). A diferença observa-se em *Tense*, em que surge (PST, None) (2/33,33%).

Os resultados relativos à distribuição das relações temporais associadas às combinações de atributos e valores dos eventos nos argumentos de razão e de resultado evidenciam algumas regularidades. As relações de sucessividade temporal (*Before* e *After*) associam-se frequentemente a eventos dinâmicos e télicos, em particular a transições, como observado por Silvano et al. (2024).

No que diz respeito às relações de inclusão (*Includes* e *Is_included*), observa-se que o intervalo temporal associado aos estados tende a incluir o evento dinâmico. Inversamente, os eventos dinâmicos podem ocorrer durante um estado, o que sugere que a natureza durativa dos estados permite que funcionem como intervalos de inclusão, justificando a presença destas relações temporais no corpus.

A relação *Simultaneous* associa-se frequentemente a situações estativas, uma vez que este tipo de relação se liga ao enquadramento do acontecimento e não à progressão narrativa ou à mudança de estado, conforme já observado por Silvano et al. (2024).

6.3. Sistematização dos resultados

Após uma análise detalhada dos resultados, e com o objetivo de proporcionar uma visão global mais clara da discussão desenvolvida, apresentam-se de forma sistemática as principais propriedades identificadas nas relações causais nas notícias da Agência Lusa escritas em português europeu. Assim, no que diz respeito à frequência, estrutura e organização das relações causais, constata-se que:

- ❖ A partir de 300 notícias, identificaram-se 829 relações causais, verificando-se que o número médio de relações por notícia é de cerca de 3.

- ❖ Em 85,77% dos casos, as relações causais estruturam-se a partir de um par de eventos associados a dois argumentos; os restantes 14,23% correspondem a casos complexos, em que um único argumento contém múltiplos eventos articulados por conjunção coordenativa copulativa ou disjuntiva.
- ❖ Quanto à ordem dos argumentos, as relações com a sequência Resultado–Razão representam 63,45%, enquanto as de Razão–Resultado correspondem a 36,55%.
- ❖ Das 829 relações causais, 79,73% são explícitas e 20,27% são implícitas.
- ❖ Os sinais linguísticos mais frequentes na marcação da causalidade explícita são: preposição (58,70%), verbo (33,28%) e conjunção (6,35%).

No que diz respeito aos eventos associados aos argumentos de razão e de resultado, das 829 relações causais extraíram-se 1804 eventos, em relação aos quais se destacam os atributos e os valores mais frequentes:

- ❖ *Class: occurrence* (78,77%)
- ❖ *Event_Type: transition* (74,19%)
- ❖ *Pos: noun* (53,33%)
- ❖ *Tense: past* (75,35%)
- ❖ *Aspect: perfective* (92,41%)
- ❖ *Vform: infinitive* (67,96%)
- ❖ *Mood: future* (66,11%)
- ❖ *Modality: poder* (62,07%)
- ❖ *Polarity: positive* (98,34%)

Embora os valores dos eventos no argumento de razão e no argumento de resultado apresentem regularidades muito semelhantes nos atributos *Class*, *Event_Type*, *Tense* e *Aspect*, observam-se diferenças relevantes nos seguintes pontos:

- ❖ Em *Pos*, no argumento de razão predomina *noun* (76,15%), enquanto no argumento de resultado predomina *verb* (59,21%).
- ❖ Em *Mood*, no argumento de razão predomina *conditional* (33,33%) e no argumento de resultado predomina *future* (66,67%).
- ❖ Em *Modality*, no argumento de razão predomina *poder* (100%) e no argumento de resultado predomina *ter_de* (50%).

Relativamente às combinações dos valores dos eventos nos argumentos de razão e de resultado, as mais frequentes correspondem às seguintes combinações por atributo:

- ❖ *Class (occurrence, occurrence)*: 65,24%
- ❖ *Event_Type (none, transition)*: 35,92%
- ❖ *Pos (noun, verb)*: 39,80%
- ❖ *Tense (none, past)*: 40,00%
- ❖ *Aspect (none, perfective)*: 45,41%
- ❖ *Vform (none, none)*: 79,49%
- ❖ *Mood (none, none)*: 97,86%
- ❖ *Modality (none, none)*: 96,84%
- ❖ *Polarity (pos, pos)*: 96,33%

Quanto à distribuição dos tipos de relações temporais no corpus (829 relações), os mais frequentes são:

- ❖ *Before*: 53,08%
- ❖ *After*: 30,16%

Ao particularizar a análise às relações temporais nas relações causais explícitas (661 ocorrências), verifica-se que os tipos mais frequentes são:

- ❖ *Before*: 55,37%
- ❖ *After*: 29,95%

Nas relações causais implícitas, os tipos mais frequentes são igualmente:

- ❖ *Before*: 44,05%
- ❖ *After*: 30,95%

Considerando a associação entre relação temporal e ordem dos argumentos, observa-se que, nas relações causais explícitas, quando a ordem é Razão–Resultado, o tipo temporal mais frequente é *After* (84,07%), e quando a ordem é Resultado–Razão, predomina *Before* (82,07%). Nas relações causais implícitas, verifica-se o mesmo padrão: com ordem Razão–Resultado predomina *After* (67,11%) e com ordem Resultado–Razão predomina *Before* (76,09%).

Quanto às combinações de valores nos atributos dos eventos associadas a relações temporais nas relações causais, e considerando os tipos *Before* e *After*, verificou-se que, de um modo geral, os valores combinados são semelhantes para ambos os argumentos, diferenciando-se sobretudo no atributo *Vform*: na relação *Before*, a combinação dominante é *None-INF* (62,86%), enquanto na relação *After* predomina *None-None* (32,08%). Quanto aos tipos *Includes* e *Is_included*, observa-se, na maioria das combinações, um comportamento de oposição: em *Includes* predominam as combinações *Class* (OCC-ST), *Event_Type* (TRAN-ST) e *Aspect* (PERF-None), enquanto em *Is_included* predominam *Class* (ST-OCC), *Event_Type* (ST-TRAN) e *Aspect* (None-PERF). Em contrapartida, regista-se semelhança em *Vform*, com predominância de (None, None). Por fim, na relação temporal *Simultaneous*, as combinações mais salientes são as que envolvem eventos estativos, nomeadamente *Class* (ST-ST) e *Event_Type* (ST-ST).

Considerações Finais

A causalidade tem sido objeto de estudo em diferentes áreas científicas devido não só ao papel central que ocupa na compreensão do que nos rodeia, como também à sua complexidade. Apesar disso, há ainda muitos problemas sobre causalidade que carecem de resposta, nomeadamente em Semântica. O objetivo principal deste percurso investigativo consistiu precisamente em compreender melhor o que define a causalidade em notícias de carácter narrativo em português europeu de modo a contribuir para uma caracterização deste tipo de relações no quadro das relações discursivas e, simultaneamente, desta dimensão das notícias narrativas.

Para alcançar esse objetivo, foram formuladas as seguintes questões de investigação:

- ❖ Q11: Quais são as principais propriedades semânticas das relações de causa em notícias de natureza narrativa em português europeu?
- ❖ Q12: Qual é a configuração da organização dos argumentos das relações causais em notícias de natureza narrativa em português europeu?
- ❖ Q13: Com que frequência ocorrem as relações causais explícitas e implícitas em notícias de agência em português europeu?
- ❖ Q14: Quais são os elementos linguísticos que contribuem para a identificação das relações de causa em notícias de natureza narrativa em português europeu?
- ❖ Q15: Entre as propriedades anotadas dos eventos representados nos argumentos das relações causais, que combinações de atributos ocorrem com maior frequência?
- ❖ Q16: Quais são as relações temporais estabelecidas entre os eventos principais dos argumentos nas relações causais em notícias de natureza narrativa em português europeu?

Foram estes os eixos que orientaram o desenvolvimento do presente trabalho em torno da causalidade em notícias em português europeu, abrangendo a relação discursiva de Causa, os elementos linguísticos que marcam esta relação, os eventos associados aos argumentos de razão e de resultado e as relações temporais que se

articulam com as relações causais. Estes aspetos foram explorados tanto no quadro concetual como na componente empírica da investigação.

No estudo realizado, confirmou-se que a causalidade constitui um conceito complexo e polissémico, objeto de múltiplas abordagens teóricas, refletindo diferentes perspetivas das várias áreas do saber. Apesar dessa diversidade concetual, mantém-se a noção fundamental de uma relação entre causa e efeito. No domínio da linguística, a causalidade é entendida semanticamente como uma relação entre um argumento de causa e um de efeito, em que o acontecimento causal desencadeia o efeito, seguindo geralmente uma lógica de precedência temporal (Barik et al., 2016; Lopes, 2004; Li, 2021; Tan et al., 2024).

A revisão da literatura apresentada no primeiro capítulo evidenciou ainda a coexistência de diferentes terminologias para designar esta relação, como causa–efeito ou agente–paciente (Soares da Silva, 1999), proposição causa–proposição efeito (Arrais, 1985), evento-causa–evento-efeito (Shibatani, 1976) e razão–resultado (ISO 24617-8, 2016). Neste estudo, optou-se pela designação razão e resultado, adotada pela Norma ISO 24617-8, por se revelar adequada à análise das relações de causa nas notícias da Agência Lusa em português europeu, ser mais transparente e ser coerente com o esquema de anotação desenhado. Para além da diversidade terminológica, identificaram-se também várias configurações das relações causais, destacando-se a causação permissiva e a causação factiva (Nedyalkov & Silnitsky, 1973), a causa de conteúdo e a causa epistémica (Sweetser, 1990; Noordman & Blijzer, 2000), bem como a causa direta e indireta (Galán Rodríguez, 1999).

A causalidade foi igualmente formalizada no domínio das relações discursivas, entendidas como as relações semânticas que contribuem para a coerência e organização de um texto. No segundo capítulo, foram analisadas várias abordagens teóricas, nomeadamente a *Theory of Discourse Coherence* (Hobbs, 1985), a *Rhetorical Structure Theory* (Mann & Thompson, 1988), a *Segmented Discourse Representation Theory* (Asher & Lascarides, 2003; Cunha & Silvano, 2009; Silvano, 2010), o *Penn Discourse Treebank* (Prasad et al., 2008) e a ISO 24617-8 (2016). Todas estas abordagens

reconhecem a causalidade como uma relação fundamental para a constituição da coerência discursiva.

O terceiro capítulo evidenciou que os elementos linguísticos envolvidos nas relações causais incluem nomes causativos, verbos causativos, advérbios, preposições e locuções prepositivas, conjunções e locuções conjuncionais, bem como construções oracionais coordenadas e subordinadas. As relações causais inferidas a partir destes elementos são consideradas explícitas, enquanto aquelas que dependem da interação entre conhecimento lexical, semântico, enciclopédico e discursivo são classificadas como implícitas.

O quarto capítulo abordou as categorias de Tempo e de Aspecto, essenciais para a interpretação das relações temporais associadas às relações causais. As relações temporais entre eventos, bem como a forma como a estrutura interna dessas situações é concetualizada, desempenham um papel central na análise da causalidade no discurso jornalístico. No âmbito das propostas teóricas de Tempo e Aspecto analisadas, destacam-se, neste contexto, as contribuições de Reichenbach (1947), Kamp & Reyle (1993) e Oliveira (2003; 2013) no domínio do Tempo, e de Vendler (1967), Moens & Steedman (1988) e Cunha (2013) no domínio do Aspecto.

Os esquemas de anotação temporal e causal analisados no quinto capítulo, nomeadamente os propostos por Girju et al. (2007), Dunietz et al. (2015; 2017), Tan et al. (2022), Pustejovsky et al. (2003b), Costa & Branco (2012), Silvano et al. (2021; 2024), Bethard et al. (2008) e Mirza & Tonelli (2014), foram fundamentais para a conceção do esquema de anotação causal desenvolvido neste trabalho.

No sexto capítulo, apresentou-se o esquema de anotação causal e temporal criado, LUSA-ACT, baseado na ISO 24617-8, ISO 24617-1 e no esquema de anotação de Silvano et al. (2021) e Leal et al. (2022), que harmoniza a representação da informação causal e temporal, permitindo uma análise mais detalhada das relações causais. O esquema proposto foi usado para anotar um corpus de 300 notícias, de natureza narrativa, redigidas em português europeu e provenientes da Agência Lusa, por nós compilado.

A análise desenvolvida permitiu responder de forma sistemática às questões de investigação formuladas, contribuindo para uma caracterização abrangente das relações causais em notícias de natureza narrativa em português europeu.

Quanto à Q11, os resultados evidenciam que as relações de causa em notícias narrativas ocorrem com alguma frequência e se caracterizam, do ponto de vista semântico, por uma estrutura nuclear composta por dois argumentos, razão e resultado, que estabelecem uma ligação entre situações predominantemente dinâmicas e télicas. Estas características inserem-se num quadro narrativo que privilegia acontecimentos terminados ou perspectivados como terminados, refletindo a função informativa das notícias e a sua orientação para a descrição de eventos do mundo.

No que diz respeito à Q12, a organização dos argumentos das relações causais revela-se flexível. Embora a configuração mais simples, com um único argumento de razão e um de resultado, seja claramente predominante, identificam-se também estruturas mais complexas, em que um dos argumentos integra múltiplas situações articuladas por mecanismos de coordenação. Observa-se ainda variação na ordem dos argumentos, sendo frequente a antecipação do resultado em relação à razão, o que sugere estratégias discursivas de destaque informativo ou de gestão da progressão temática.

Quanto à Q13, verifica-se que as relações causais ocorrem maioritariamente de forma explícita no discurso noticioso, embora as relações implícitas também estejam presentes de forma consistente. Este padrão confirma a tendência do género jornalístico para privilegiar a explicitação das relações semânticas, facilitando a interpretação do leitor, sem excluir, contudo, mecanismos inferenciais que permitem a recuperação da causalidade a partir do contexto.

No âmbito da Q14, a análise dos mecanismos linguísticos envolvidos na identificação das relações de causa mostra uma grande diversidade de recursos. Destacam-se, em particular, as preposições e locuções prepositivas e os verbos causativos, que desempenham um papel central na marcação explícita da causalidade. As conjunções surgem com menor frequência, enquanto nomes, adjetivos e advérbios assumem um papel mais marginal.

Relativamente à Q15, a análise das propriedades dos eventos associados aos argumentos de razão e de resultado evidencia combinações de atributos recorrentes. Os eventos de razão tendem a apresentar realizações nominais, enquanto os eventos de resultado são maioritariamente verbais. De forma geral, as combinações mais frequentes confirmam a predominância de eventos dinâmicos, télicos e perfetivos, expressos em contextos de polaridade afirmativa.

Por fim, no que concerne à Q16, as propriedades temporais e aspetuais dos eventos permitem justificar relações temporais sistemáticas entre os argumentos causais. A sucessividade temporal emerge como o padrão dominante, refletindo a lógica narrativa de encadeamento dos acontecimentos, embora também se observem relações de inclusão temporal e de simultaneidade. Estes padrões mantêm-se de forma consistente independentemente de a relação causal ser explícita ou implícita e da ordem em que os argumentos são apresentados.

Este trabalho apresenta três contributos principais para o estudo das relações causais. O primeiro consiste no aprofundamento da análise semântica da causalidade, ao clarificar propriedades fundamentais para a caracterização semântica da relação de Causa e para os mecanismos envolvidos na sua inferência. O segundo contributo reside na proposta do esquema de anotação LUSA-ACT, que articula de forma integrada as dimensões causal e temporal, promovendo uma abordagem mais consistente da anotação semântica. Por fim, a investigação deu origem a um corpus anotado com relações causais e temporais em português europeu, a disponibilizar em acesso aberto, constituindo um recurso valioso para a investigação linguística e computacional e respondendo à escassez de recursos deste tipo para esta variedade do português.

Concluída a investigação, reconhecem-se algumas limitações, nomeadamente a não exploração aprofundada de relações causais complexas em que uma relação integra ou engloba outras relações causais, bem como de estruturas em que um mesmo argumento desempenha simultaneamente papéis distintos em diferentes relações causais. Estas limitações abrem caminho para investigações futuras, que poderão aprofundar a análise de relações causais complexas e de encadeamentos causais mais elaborados em notícias.

Como investigação futura, é também nosso objetivo integrar mais relações discursivas no esquema LUSA-ACT, de modo a permitir representar outras relações de sentido construídas em notícias.

Referências Bibliográficas

- Abreu, C. C. S. R., Pezatti, E. G., & Novaes-Marques, N. B. (2016). As orações causais. In E. G. Pezatti (Org.), *Construções subordinadas na lusofonia: Uma abordagem discursivo-funcional* (pp. 121–151). Editora UNESP. <https://doi.org/10.7476/9788568334805>
- Akoff, G. (1970). *Irregularity in syntax*. Holt, Rinehart and Winston.
- Alexiadou, A. (2001). *Functional structure in nominals: Nominalization and ergativity*. Amsterdam: John Benjamins.
- Ali, M. S. (1971). *Gramática histórica da língua portuguesa* (7.^a ed.). Rio de Janeiro: Livraria Acadêmica. (Original publicado em 1931).
- Amorim, E., Leal, A., Yu, N., Silvano, P. M., & Jorge, A. M. (2025). Visual representations of temporal relations between events and time expressions in news stories. In *Proceedings of the 19th Linguistic Annotation Workshop (LAW-XIX-2025)* (pp. 250–263). Vienna, Austria: Association for Computational Linguistics.
- Armstrong, D. M. (1997). *A world of states of affairs*. Cambridge University Press.
- Arrais, T. C. (1985). As construções causativas em português. *Alfa*, 29, 41–58.
- Asher, N., & Lascarides, A. (2003). *Logics of conversation*. Cambridge: Cambridge University Press.
- Bach, E. (1986). The algebra of events, *Linguistics and Philosophy*, 9, pp. 5-16
- Barik, B., Marsi, E., & Öztürk, P. (2016). Event causality extraction from natural science literature. *Research in Computing Science*, 117, 97–107.
- Bartalesi Lenzi, V., Sprugnoli, R., Tonelli, S., & Lenci, A. (2012). Annotating causal relations in the TempEval-3 corpus. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012)* (pp. 1787–1794). European Language Resources Association (ELRA).
- Bazzanella, C. (1995). I segnali discorsivi. In L. Renzi, G. Salvi, & A. Cardinaletti (Eds.), *Grande Grammatica Italiana di Consultazione* (Vol. 3, pp. 225–257). Bologna: Il Mulino.
- Bennett, J. F. (1988). *Events and Their Names*. Indianapolis, IN: Hackett Publishing.

- Bethard, S., & Martin, J. H. (2008). Learning semantic links from a corpus of parallel temporal and causal relations. In *Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics (ACL 2008)* (pp. 177–180). Association for Computational Linguistics.
- Blakemore, D. (1987). *Semantic Constraints on Relevance*. Oxford: Blackwell.
- Borer, H. (2003). Exo-skeletal vs. endo-skeletal explanations: Syntactic projections and the lexicon. In J. Moore & M. Polinsky (Eds.), *The nature of explanation in linguistic theory* (pp. 31–67). Stanford, CA: CSLI Publications.
- Brito, A. M. (2003). Aspectos Sintáticos da Gramática do Português. In M. H. M. Mateus et al. (Eds.), *Gramática da Língua Portuguesa* (5ª ed., pp. 127–178). Lisboa: Caminho.
- Brito, A. M., & Oliveira, F. (1997). Nominalization, aspect and argument structure. In G. Matos et al. (Eds.), *Interfaces in linguistic theory: Selected papers from the International Conference on Interfaces in Linguistics, Porto, November 13–17, 1995* (pp. 57–80). Associação Portuguesa de Linguística/Edições Colibri.
- Campbell, K. (1990). *Abstract Particulars*. Oxford: Basil Blackwell.
- Carlson, G. N. (1977). *Reference to kinds in English* (Unpublished doctoral dissertation). University of Massachusetts Amherst.
- Caselli, T., Russo, I., & Sprugnoli, R. (2011). It-TimeML: An Italian temporal annotation standard. In *Proceedings of the Fifth Linguistic Annotation Workshop* (pp. 64–72). Association for Computational Linguistics.
- Castilho, A. T. de. (2014). *Nova gramática do português brasileiro* (1ª ed., 3ª. reimpressão). São Paulo, SP: Editora Contexto.
- Costa, A. P., Moreira, A., & Sá, P. (Orgs.). (2021). *Reflexões em torno de metodologias d -e investigação: Análise de dados* (Vol. 3). UA Editora. <https://doi.org/10.34624/dws9-6j98>
- Costa, F. (2012). *Processing temporal information in unstructured documents* (Doctoral dissertation, Faculdade de Ciências da Universidade de Lisboa).

- Costa, F., & Branco, A. (2012). TimeBankPT: A TimeML annotated corpus of Portuguese. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC 2012)*. European Language Resources Association (ELRA).
- Couper-Kuhlen, E., & Kortmann, B. (2000). Introduction. In E. Couper-Kuhlen & B. Kortmann (Eds.), *Cause, condition, concession, contrast: Cognitive and discourse perspectives* (pp. 1–10). Mouton de Gruyter.
- Cresswell, M. J. (1981). Adverbs of causation. In H. J. Eikmeyer & H. Rieser (Eds.), *Words, worlds, and contexts: New approaches in word semantics* (pp. 21–37). De Gruyter. <https://doi.org/10.1515/9783110842524-003>
- Crothers, E. (1979). *Paragraph structure inference*. Norwood, NJ: Ablex Publishing Corporation.
- Cunha, C. F., & Cintra, L. (1985). *Nova gramática do português contemporâneo* (2.ed.). Rio de Janeiro: Nova Fronteira.
- Cunha, L. F. (2004). *Semântica das Predicações estativas. Para uma Caracterização Aspectual dos Estados* (Tese de Doutorado, Universidade do Porto).
- Cunha, L. F. and Silvano, P. (2009). “O papel das restrições aspectuais nas relações retóricas: o caso das frases complexas com quando”. In *Textos Seleccionados do XXIV Encontro Nacional da Associação Portuguesa de Linguística* (pp. 239-250). Lisboa: APL.
- Cunha, L. F. (2013). Aspeto. In E. B. P. Raposo, M. F. B. do Nascimento, M. A. C. da Mota, L. Segura, & A. Mendes (Orgs.), *Gramática do português* (Vol. 1, pp. 585-616). Fundação Calouste Gulbenkian.
- Cunha, L. F., Leal, A., & Silvano, M. D. P. (2015). Relações retóricas e temporais em construções gerundivas adverbiais. *Estudos de semântica*.
- Davidson, D. (1967). Causal relations. *The Journal of Philosophy*, Vol. 64, No. 21, *Sixty-Fourth Annual Meeting of the American Philosophical Association, Eastern Division*. (Nov. 9, 1967), pp. 691-703.
- Dik, S. C. (1997). *The theory of functional grammar. Part I: The structure of the clause* (2nd rev. ed.). Mouton de Gruyter.

- Dik, S. C., Hengeveld, K., Vester, E., & Vet, C. (1990). *The hierarchical structure of the clause and typology of adverbial satellites*. In J. Nuyts, A. M. Bolkestein, & C. Vet (Eds.), *Layers and levels of representation in language theory: A functional view* (pp. 25–70). John Benjamins.
- Dirven, R. (1995). The construal of cause: The case of cause prepositions. In J. R. Taylor & R. E. MacLaury (Eds.), *Language and the cognitive construal of the world* (pp. 95–118). Berlin: Mouton de Gruyter.
- Dowty, D. R. (1979). *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ* (Vol. 7). Springer.
- Dretske, F. (1970). Referring to events. *Midwest Studies in Philosophy*, 2(1), 90–99.
- Dunietz, J., Levin, L., & Carbonell, J. (2015). Annotating causal language using corpus lexicography of constructions. In *Proceedings of the 9th Linguistic Annotation Workshop (LAW IX)* (pp. 188–196). Association for Computational Linguistics.
- Dunietz, J., Levin, L., & Carbonell, J. (2017). The BECauSE Corpus 2.0: Annotating causality and overlapping relations. In *Proceedings of the 11th Linguistic Annotation Workshop* (pp. 95–104). Association for Computational Linguistics.
- Fernandes, A. L., Silvano, P., Leal, A., Guimarães, N., Rb-Silva, R., Cunha, L. F., & Jorge, A. (2025a). The incremental process of building an annotation scheme for clinical narratives in Portuguese: The contribution of human variation analysis. In *Proceedings of the 19th Linguistic Annotation Workshop (LAW-XIX-2025)* (pp. 332–343). Vienna, Austria: Association for Computational Linguistics.
- Fernandes, A. L., Silvano, P., Guimarães, N., Rb-Silva, R., Munna, T. A., Cunha, L. F., Leal, A., Campos, R., & Jorge, A. M. (2025b). Human Experts vs. Large Language Models: Evaluating Annotation Scheme and Guidelines Development for Clinical Narratives. *Proceedings of Text2Story — Eighth Workshop on Narrative Extraction From Texts* (CEUR Vol-3964, pp. 149–160). <https://ceur-ws.org/Vol-3964/paper13.pdf>
- Fillmore, C. J. (1974). Pragmatics and the description of discourse. In C. Fillmore, G. Lakoff, & R. Lakoff (Eds.), *Berkeley studies in syntax and semantics* (pp. 143–166). University of California.

- Fraser, B. (1990). An approach to discourse markers. *Journal of Pragmatics*.
[https://doi.org/10.1016/0378-2166\(90\)90096-V](https://doi.org/10.1016/0378-2166(90)90096-V)
- Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics*, 31(7), 931–952.
- Galán Rodríguez, C. (1999). La subordinación final y causal. En I. Bosque & V. Demonte. (Eds.), *Entre la oración y el discurso. Morfología* (Vol. 3, cap. 56, pp. 3597–3642). Madrid: Editorial Espasa Calpe, S.A.
- García Fernández, L. (1999). Los complementos adverbiales temporales. La subordinación temporal. En I. Bosque & V. Demonte (Eds.), *Las construcciones sintácticas fundamentales. Relaciones temporales, aspectuales y modales* (Vol. 2, cap. 48, pp. 3129–3208). Madrid: Editorial Espasa Calpe, S.A.
- Girju, R. (2003). Automatic Detection of Causal Relations for Question Answering. *Proceedings of the ACL 2003 Workshop on Multilingual Summarization and Question Answering* (pp. 76–83). <https://doi.org/10.3115/1119312.1119322>
- Girju, R., Nakov, P., Nastase, V., Szpakowicz, S., Turney, P., & Yuret, D. (2007). SemEval 2007 task 04: Classification of semantic relations between nominals. *Proceedings of the 4th International Workshop on Semantic Evaluation (SemEval-2007)* (pp. 13–18). Association for Computational Linguistics.
- Givón, T. (1975). Cause and control: On the semantics of interpersonal manipulation. In J. Kimball (Ed.), *Syntax and semantics* (Vol. 4, pp. 59–89). Academic Press.
- Gonçalves, D. (2019). *O livro de estilo no jornalismo de agência: o caso da Lusa* (Dissertação de Mestrado). Universidade do Porto.
- Grimes, J. (1975). *The thread of discourse*. Mouton.
- Grimshaw, J. (1990). *Argument structure*. Cambridge, Massachusetts, The MIT Press.
- Hitchcock, C. R. (1998). Causal knowledge: That great guide of human life. *Communication & Cognition*, 31(4), 271-296.
- Hobbs, J. R. (1985). *On the coherence and structure of discourse* (Report No. CSLI-85-37). Center for the Study of Language and Information.
- Hume, D. (1965). An abstract of a treatise of human nature. Hamden, Connecticut: Archon Books. (Original work published 1740).

- Hume, D. (n.d.). *Investigação acerca do entendimento humano* (A. Aiex, Trad.). Recuperado de <http://br.egroups.com/group/acropolis/>
- Im, S., Lee, K., & Pustejovsky, J. (2009). KTimeML: Specification of temporal and event expressions in Korean text. *Proceedings of the ACL-IJCNLP 2009 Workshop on Time and Event Processing* (pp. 49–56). Association for Computational Linguistics.
- ISO 24617-1. (2012). Language resource management - Semantic annotation framework (SemAF) - Part 1: Time and events (SemAF-Time, ISO-TimeML), Standard, Geneva, CH, 2012.
- ISO 24617-4. (2014). Language resource management- Semantic annotation framework (SemAF) - Part 4: Semantic roles (SemAF-SR).
- ISO 24617-7. (2019). Language resource management Spatial information (SemAF) – Part 7: Reference annotation framework (ISO-Space).
- ISO 24617-8. (2016). Language resource management - Part 8: Semantic relations in discourse (DR-Core). Geneva, Switzerland: ISO.
- ISO 24617-9. (2019) Language resource management Semantic annotation framework (SemAF) - Part 9: Reference annotation framework (RAF).
- Kamp, H., & Reyle, U. (1993). *From discourse to logic: Introduction to model-theoretic semantics of natural language, formal logic and discourse representation theory*. Kluwer Academic Publishers.
- Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.
- Khoo, C., Chan, S., & Niu, Y. (2002). The many facets of the cause–effect relation. In R. Green, C. A. Bean, & S. H. Myaeng (Eds.), *The semantics of relationships: An interdisciplinary perspective* (pp. 51–70). Kluwer Academic Publishers.
- Knott, A. and R. Dale (1994). Using linguistic phenomena to motivate a set of coherence relations. *Discourse Processes*, 18, pp. 35–62.
- Kury, A. M. (1985). *Subordinação e coordenação no sistema das orações do português*. Tempo Brasileiro.
- Lakoff, G. (1970). *Irregularity in syntax*. Holt, Rinehart and Winston.
- Lapesa, R. (1978). *Historia de la lengua española* (9.ª ed.). Madrid: Gredos.

- Leal, A. (2009). *Semântica Aspectual e Nominal—Contributo das Expressões Nominais. P-ara a Construção Aspectual das Frases* (Faculdade de Letras da Universidade do Porto Tese de doutoramento). <https://scholar.google.com/scholar?cluster=12051810427835546918>
- Leal, A., Silvano, P., Amorim, E., Cantante, I., Silva, F., Jorge, A. M., & Campos, R. (2022, June). The place of ISO-Space in Text2Story multilayer annotation scheme. In *Proceedings of the 18th Joint ACL-ISO Workshop on Interoperable Semantic Annotation within LREC2022* (pp. 61-70).
- Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago Press.
- Li, J. (2021). The semantics of causative events: From the perspective of event integration [因果关系事件语义学：事件融合视角]. Jinan University Press.
- Lobo, M. (2003). *Aspectos da sintaxe das orações subordinadas adverbiais* (Tese de doutoramento). Lisboa: Universidade Nova de Lisboa.
- Longacre, R. (1976). *An anatomy of speech notions*. The Peter de Ridder Press.
- Lopes, M. H. C. C. (2004). *Aspectos sintácticos, semânticos e pragmáticos das construções causais: Contributo para uma reflexão sobre o ensino da gramática* (Tese de doutoramento, Universidade do Porto). Faculdade de Letras da Universidade do Porto.
- Lopes, A. C., & Carrilho, E. (2013). *Texto, coesão e coerência*. Almedina.
- Lopes, A. C. M., & Carrilho, E. (2020). Discurso e marcadores discursivos. In E. B. P. Raposo, M. F. B. do Nascimento, M. A. C. da Mota, L. Segura, A. Mendes, & A. Andrade (Orgs.), *Gramática do português* (Vol. III, pp. 2667–2697). Fundação Calouste Gulbenkian.
- Lusa – Agência de Notícias de Portugal. (2019). *Livro de estilo da Lusa* (A. Folhas, F. Peixeiro, J. R. Simões, P. Agostinho, N. Simas, P. A. Nogueira, P. Albuquerque, R. J. Pinto, & T. Petinga, Revs.). Lusa – Agência de Notícias de Portugal.
- Mackie, J. L. (1980). *The cement of the universe: A study of causation*. Oxford University Press.

- Mann, W. C., & Thompson, S. A. (1987). *Rhetorical Structure Theory: A theory of text organisation* (Technical Report ISI/RS-87-190). Information Sciences Institute.
- Marcos Marín, F. (1979). *Sintaxis del español actual: La oración simple*. Madrid: Gredos.
- Martín Zorraquino, M.^a A., & Portolés Lázaro, J. (1999). Los marcadores del discurso. In I. Bosque & V. Demonte (Eds.), *Gramática descriptiva de la lengua española* (Vol. 3, pp. 4051–4213). Madrid: Espasa Calpe.
- McCawley, J. D. (1968). Lexical insertion in a transformational grammar without deep structure. In B. Darden et al. (Eds.), *Papers from the Fourth Regional Meeting of the Chicago Linguistics Society* (pp. 71–80). University of Chicago.
- Mellor, D. H. (1980). On things and causes in spacetime. *The British Journal for the Philosophy of Science*, 31(3), 282–288.
- Mendes, A., Lopes, J., & Silva, J. (2023). Using a discourse bank and a lexicon for the automatic identification of discourse connectives. *Proceedings of the International Conference on Computational Linguistics (COLING 2023)*, 1–12.
- Mendes, A., Lopes, J., & Silva, J. (2024). LDM-PT: A lexicon of discourse markers for Portuguese. *Proceedings of the Language Resources and Evaluation Conference (LREC 2024)*, 1–12.
- Menzies, P. (1989). A unified account of causal relations. *Australasian Journal of Philosophy*, 67, pp. 59-83.
- Meyer, P. G. (2000). The relevance of causality. In E. Couper-Kuhlen & B. Kortmann (Eds.), *Cause, condition, concession, contrast: Cognitive and discourse perspectives* (Vol. 33, pp. 9-34). Mouton de Gruyter.
- Mill, J. S. (1886). *A system of logic, ratiocinative and inductive: Being a connected view of the principles of evidence and the methods of scientific investigation*. Longmans, Green, and Co., London.
- Mirza, P., Sprugnoli, R., Tonelli, S., & Speranza, M. (2014). Annotating causality in the TempEval-3 corpus. *Proceedings of the EACL 2014 Workshop on Computational Approaches to Causality in Language* (pp. 10–19). Association for Computational Linguistics.

- Moens, M. (1987). *Tense, Aspect and Temporal Reference* (Ph.D. thesis). Edinburgh: Centre for Cognitive Science, University of Edinburgh, Scotland.
- Moens, M. & M. Steedman (1987). "Temporal Ontology and Temporal Reference". In *Computational Linguistics*, Volume 14, number 2, 15-28.
- Móia, T., & Alves, A. T. (2003). Tempo e aspeto. In M. H. M. Mateus, A. M. Brito, I. Duarte, & I. H. Faria (Eds.), *Gramática da Língua Portuguesa* (2.^a ed., pp. 409–462). Lisboa: Caminho.
- Nedjalkov, V. P., & Silnitsky, G. G. (1973). The typology of morphological and lexical causatives. In F. Kiefer (Ed.), *Trends in Soviet Theoretical Linguistics* (pp. 1–32). Dordrecht: Reidel.
- Neves, M. H. M. (2000). *Gramática de usos do português*. Unesp.
- Noordman, L. G. M., & de Blijzer, F. (2000). On the processing of causal relations. In E. Couper-Kuhlen & B. Kortmann (Eds.), *Cause, condition, concession, contrast: Cognitive and discourse perspectives* (Vol. 33, pp. 9–34). Mouton de Gruyter.
- Nunes, S., Jorge, A., Leal, A., Amorim, E., Sousa, H., Cantante, I., Silvano, P., & Campos, R. (2023). *Text2Story Lusa* [Data set]. INESC TEC. <https://doi.org/10.25747/ET95-BX90>
- Nunes, S., Jorge, A. M., Amorim, E., Sousa, H., Leal, A., Silvano, P., Cantante, I., & Campos, R. (2024). Text2Story Lusa: A Dataset for Narrative Analysis in European Portuguese News Articles. *International Conference on Language Resources and Evaluation*.
- Nuss, V. M. (2021). *As relações de causalidade por meio do sintagma preposicionado por + substantivo* (Tese de doutorado, Universidade Estadual Paulista "Júlio de Mesquita Filho", São José do Rio Preto).
- Oliveira, F. (2003). Tempo e aspeto. In M. H. M. Mateus, A. M. Brito, I. Duarte, & I. H.
- Oliveira, F. (2013). Tempo verbal. In E. B. P. Raposo, M. F. B. do Nascimento, M. A. C. da. Mota, L. Segura, & A. Mendes (Orgs.), *Gramática do português* (Vol. 1, pp. 509-546). Fundação Calouste Gulbenkian.
- Oversteegen, L. (2005). Causality and Tense—Two Temporal Structure Builders. *Journal of Semantics*, 22(3), 307–337. <https://doi.org/10.1093/jos/ffh021>

- Paul, L. A. (2000). *Aspect causation*. *The Journal of Philosophy*, 97, pp. 223–34.
- Piskorski, J., Jorge, A., Silvano, M. D. P., Guimarães, N., Pacheco, A. F., & Yu, N. (2024). *Overview of the CLEF-2024 CheckThat! Lab task 3 on persuasion techniques*. In Proceedings of the CLEF 2024 Working Notes.
- Portolés, J. (1993). La distinción entre los conectores y otros marcadores del discurso en español. *Verba: Anuario Galego de Filoloxía*, 20, 141–170.
- Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A., & Webber, B. (2008). *The Penn Discourse TreeBank 2.0*. Philadelphia, PA : University of Pennsylvania.
- Prasad, R., Joshi, A., & Webber, B. (2010). Realization of discourse relations by other means: Alternative lexicalizations. In *Coling 2010: Poster Volume* (pp. 1023–1031). Beijing.
- Prasad, R., Webber, B., Dinesh, N., Lee, A., & Joshi, A. (2019). *The Penn Discourse Treebank 3.0*. Philadelphia: Linguistic Data Consortium.
- Pustejovsky, J. (1991). The generative lexicon. *Computational linguistics*, 17(4), 409–441.
- Pustejovsky, J., Hanks, P., Saurí, R., See, A., Gaizauskas, R., Setzer, A., Sundheim, B., Radev, D., Day, D., Ferro, L., & Lazo, M. (2003a). The TimeBank corpus. *Proceedings of the Corpus Linguistics 2003 Conference*, 647–656.
- Pustejovsky, J., Castaño, J., Ingria, R., Sauri, R., Gaizauskas, R., Setzer, A., & Katz, G. (2003b). TimeML: Robust specification of event and temporal expressions in text. *New Directions in Question Answering*, 28–34.
- Pustejovsky, J., Lee, K., Bunt, H., & Romary, L. (2010). ISO-TimeML: An international standard for semantic annotation. *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC 2010)*. European Language Resources Association (ELRA).
- Pustejovsky, J., & Stubbs, A. (2012). *Natural language annotation for machine learning*. O'Reilly Media.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. Harlow, England: Longman.

- Raposo, E. B. P. (2013). Estrutura da frase. In E. B. P. Raposo, M. F. B. do Nascimento, M. A. C. da Mota, L. Segura, & A. Mendes, & A. Andrade (Orgs.), *Gramática do português* (Vol. I, pp. 303–400). Lisboa: Fundação Calouste Gulbenkian.
- Raposo, E. B. P., & Xavier, M. F. (2013). Preposição e sintagma preposicional. In E. B. P. Raposo, M. F. B. do Nascimento, M. A. C. da Mota, L. Segura, A. Mendes & A. Andrade (Orgs.), *Gramática do português* (Vol. II, pp. 1497–1566). Lisboa: Fundação Calouste Gulbenkian.
- Reinchenbach, H. (1947). *Elements of Symbolic Logic*. London: Macmillan.
- Rink, B., Bejan, C. A., & Harabagiu, S. (2010). Learning textual graph patterns to detect causal event relations. *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI 2010)*. Association for the Advancement of Artificial Intelligence.
- Rocha Lima, C. H. (2011). *Gramática normativa da língua portuguesa* (49.^a ed.). Rio de Janeiro: José Olympio.
- Schiffrin, D. (1987). *Discourse Markers*. Cambridge: Cambridge University Press.
- Schorup, L. (1985). *Common discourse particles in English conversation: like, well, y'know*. New York: Garland.
- Shibatani, M. (1976). The grammar of causative constructions: A conspectus. In M. Shibatani (Ed.), *Syntax and semantics* (Vol. 6, The grammar of causative constructions, pp. 1–40). Academic Press.
- Stenetorp, P., Pyysalo, S., Topić, G., Ohta, T., Ananiadou, S., & Tsujii, J. (2012). *BRAT: A web-based tool for NLP-assisted text annotation*. In *Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 102–107). Association for Computational Linguistics.
- da Silva, P. N. (2005). *O tempo no texto: Contributos para o estudo da expressão do tempo em sequências textuais* (Tese de doutoramento, Universidade Aberta). Universidade Aberta.
- Silvano, P. (2002). Sobre a semântica da sequência de tempos em Português Europeu. *Análise das relações temporais em frases complexas com completivas*.

- Silvano, P. (2010). *Temporal and rhetorical relations: The semantics of sentences with adverbial subordination in European Portuguese* (Doctoral dissertation). Faculdade de Letras da Universidade do Porto, Universidade do Porto.
- Silvano, P., Leal, A., Silva, F., Cantante, I., Oliveira, F., & Jorge, A. M. (2021). Developing a multilayer semantic annotation scheme based on ISO standards for the visualization of a newswire corpus. In *Proceedings of the 17th Joint ACL-ISO Workshop on Interoperable Semantic Annotation* (pp. 1-13).
- Silvano, M. D. P., Amorim, E., Leal, A., Cantante, I., Jorge, A., Campos, R., & Yu, N. (2024). Untangling a web of temporal relations in news articles. In *Proceedings of Text2Story 2024-seventh workshop on narrative extraction from texts*.
- Soares da Silva, A. (1999). *A semântica de deixar: Uma contribuição para a abordagem cognitiva em semântica lexical*. Fundação Calouste Gulbenkian.
- Sousa, J. P. (2002). Por que as notícias são como são? Construindo uma teoria da notícia. *Biblioteca On-line de Ciências da Comunicação*, 01-17.
- Strawson, P. F. (1985). Causation and explanation. In B. Vermazen & M. B. Hintikka (Eds.), *Essays on Davidson: Actions and events* (pp. 115–135). Clarendon Press.
- Sweetser, E. (1990). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge University Press.
- Szeto, Y.-K. (1988). The semantics of causative and agentive verbs. *Cahiers Linguistiques d'Ottawa*, 16, 1-51.
- Talmy, L. (1976). Semantic causative types. In M. Shibatani (Ed.), *Syntax and semantics: Vol. 6. The grammar of causative constructions* (pp. 43–116). Academic Press.
- Talmy, L. (1985). *Lexicalization patterns: Semantic structure in lexical forms*. In T. Shopen (Ed.), 57-149.
- Talmy, L. (1988). Force dynamics in language and cognition. *Cognitive Science*, 12(1), 49–100.
- Tan, F. A., Hürriyetoğlu, A., Caselli, T., Oostdijk, N., Nomoto, T., Hettiarachchi, H., Ameer, I., Uca, O., Liza, F. F., & Hu, T. (2022). The Causal News Corpus: Annotating causal relations in event sentences from news. *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022)*, 2298–2310.

- Tamura, S., Hara, Y., Kim, Y., & Sakai, H. (2010). *Types of causal relations: A survey*. *Kyoto University Linguistic Research*, 29, 153–170.
- Thompson, J. J. (1987). Verbs of action. *Synthese*, 72(1), 103-122.
- Tomaszewska, A., Silvano, P., Leal, A., & Amorim, E. (2024). ISO 24617-8 applied: Insights from multilingual discourse relations annotation in English, Polish, and Portuguese. In H. Bunt et al. (Eds.), *Proceedings of the 20th Joint ACL - ISO Workshop on Interoperable Semantic Annotation @ LREC-COLING 2024* (pp. 99–110).
- Van den Broek, P. (1989). Causal reasoning and inference making in judging the importance of story statements. *Child development*, 286-297.
- Vendler, Z. (1967). Causal relations. *The Journal of Philosophy*, 64 (21), 704–713.
- Vendler, Z. (1967). Verbs and Times, in *Linguistics in Philosophy*, 4, Cornell University Press: Ithaca, New York, 97-121.
- Wolff, P. (2007). Representing causation. *Journal of Experimental Psychology: General*, 136(1), 82–111.
- Yu, N. (2024). Leituras semânticas de preposições de origem *desde* e *a partir de* em frases em português europeu. *ElingUP: Revista Eletrónica de Linguística dos Estudantes da Universidade do Porto*, 12 (1). <https://ojs.letras.up.pt/index.php/elingUP/article/view/13824>