

A estrutura temporal do lead em notícias de agência em Português Europeu

Ana Filipa Pacheco¹

¹ Faculdade de Letras da Universidade do Porto

Resumo

O presente artigo pretende explorar as relações temporais que se estabelecem entre as situações representadas em notícias de agência, em português europeu. Para tal, recorreu-se a um *corpus* composto por trinta notícias, cujos *leads* foram anotados com recurso à ferramenta BRAT (brat rapid annotation tool). O processo de anotação seguiu a proposta de algumas das etiquetas e atributos do esquema de anotação proposto por Silvano et al. (2021), desenvolvido com base na norma ISO 24617-1:2012 (SemAF-Time, ISO-TimeML). Os resultados encontrados mostram que o *lead* se caracteriza por apresentar situações representadas, predominantemente, por nomes e verbos; os verbos encontram-se, sobretudo, no pretérito perfeito do indicativo. As relações temporais mais frequentes dizem respeito a relações de anterioridade, inclusão, identidade e simultaneidade, não se verificando sucessividade temporal.

Palavras-chave

Estrutura temporal, relações temporais entre situações, anotação, *corpus* de notícias

Abstract

The aim of this paper is to explore the temporal relations established between the situations represented in agency news stories in European Portuguese. To achieve this, we used a corpus made up of thirty news items, whose leads were annotated using the BRAT tool (brat rapid annotation tool). The annotation process followed the proposal of some of the labels and attributes of the annotation scheme presented by Silvano et al. (2021), developed using the ISO 24617-1:2012 standard (SemAF-Time, ISO-TimeML). The results show that the lead is characterised by situations represented predominantly by nouns and verbs; the verbs are mainly in the past perfect tense. The most frequent temporal relations are those of anteriority, inclusion, identity, and simultaneity, but no temporal succession.

Keywords

Temporal structure, temporal relations between situations, annotation, news corpus

1. Introdução

A análise linguística dos elementos que constituem as narrativas tem fomentado um interesse crescente na comunidade científica. As notícias de natureza predominantemente narrativa constituem uma fonte rica de informação, pela sua organização e apresentação particulares, quando em comparação com as narrativas de outros géneros textuais. Este tipo de notícias fornece dados cuja análise permite explorar, por exemplo, a forma como as situações que nelas são relatadas se organizam e apresentam face à sua cronologia e relevância temporal.

Bell (1995) defende que as notícias incluem um tipo particular de narrativa em que se verifica uma ordem não cronológica de situações, que são apresentadas tendo em consideração uma pirâmide invertida: as situações mais relevantes são introduzidas e posteriormente retomadas, uma e outra vez, em maior detalhe. Esta especificidade tem levantado questões (cf. Bell, 1995: 319) sobre a organização temporal das notícias e a forma como as situações se estruturam dentro das mesmas. Por outro lado,

¹ anafilipaspacheco@gmail.com
ORCID 0009-0003-6135-8479

a escassez de estudos no que respeita à estrutura temporal das notícias de agência, tal como Silvano et al. (2021) constata, expõe a necessidade de aprofundamento linguístico a este nível.

A linguística e, particularmente, a semântica e a análise da estrutura temporal podem contribuir para um melhor entendimento das narrativas do género notícia.

Propõe-se, por isso, com este artigo, analisar o modo como os *leads* das notícias se organizam temporalmente, ao proceder-se à averiguação das relações temporais mais frequentes entre as situações do *lead* de notícias de agência em português europeu. Os objetivos específicos focam-se em: por um lado, (i) identificar as situações do *lead* de notícias e algumas das suas propriedades aspetuais; e, por outro, (ii) identificar as relações temporais que se estabelecem entre essas situações, determinando-se, no processo, as relações mais frequentes e analisando-se a presença (ou ausência) de sucessividade temporal.

Nesse sentido, centrar-nos-emos nas seguintes questões de investigação:

- i. Que tipo de situações caracterizam o *lead* de notícias de agência?
- ii. Que tipo de relações temporais se estabelecem no *lead* de notícias de agência?
- iii. Quais as relações temporais mais frequentes no *lead* de notícias de agência?
- iv. Há sucessividade temporal no *lead* de notícias de agência?

Ao longo das próximas páginas, será feita a contextualização teórica sobre o tema, que começa com uma breve revisão sobre notícias e as suas características (ponto 2), e continua com algumas questões sobre tempo e aspeto (ponto 3) e sobre anotação da informação temporal (ponto 4). Segue-se a apresentação do estudo desenvolvido, na parte cinco, através da qual far-se-á a descrição do *corpus* utilizado e descrever-se-ão o esquema e o processo de anotação desenvolvidos. No ponto 6, apresentar-se-ão detalhadamente os resultados do estudo, assim como a sua análise e discussão. Por fim, o presente trabalho culmina com as considerações finais, em que se concluem sucintamente os pontos levantados e abordados no decurso deste artigo (cf. ponto 7).

2. Notícias: algumas características

Os textos de van Dijk (e.o. 1983, 1985, 1988) assumem uma grande importância para a compreensão do género notícia. Um dos seus contributos teóricos faz menção à superestrutura em que a notícia se encontra organizada: (i) o sumário/introdução, constituído pelas partes título e *lead*; (ii) os episódios, que incluem as situações e as reações; e (iii) os comentários, que são tecidos face aos dois anteriores e estão relacionados com as crenças do jornalista/jornal. Por outras palavras, a notícia apresenta um título e um *lead*, seguidos de um corpo. O *lead* corresponde ao primeiro parágrafo da notícia, que visa captar a atenção do leitor e resumir os principais aspetos da notícia de forma sucinta e relevante (Lusa, 2019), e procura responder às perguntas quem, o quê, quando, onde, como e porquê – os 5W1H (“who, what, when, where, why, how”) (e.o. Carnaz et al., 2021; Público, 2005), seguindo-se uma ordem decrescente de importância de acordo com a qual a informação é apresentada e organizada – a chamada pirâmide invertida (e.o. Bell, 1995; Bandler et al., 2018).

A pirâmide invertida é, assim, um estilo jornalístico de escrita em que a informação mais importante é apresentada no início, seguida de outros elementos de importância decrescente. Bell (1991) salienta o modo como esta estrutura ajuda a captar rapidamente a atenção do leitor, ao apresentar as informações mais significativas no início da notícia. Este método ajuda os leitores a apreenderem os factos essenciais – os tópicos discursivos – mesmo que só leiam o início do texto.

Bell (e.o. 1991, 1995, 2006) menciona, ainda, que a especificidade inerente às narrativas presentes nas notícias fornece uma base para o entendimento da importância do estudo linguístico no contexto dos *media*. No seguimento, o autor menciona que o parágrafo *lead* atua na notícia como um resumo,

sumarizando a ação central e estabelecendo o âmbito do que está a ser contado, e assumindo-se como o seu personagem principal.

Nas palavras de Silvano et al. (2021: 2), “news may frequently assume the format of a story that reports on current events² involving one or more entities in given time and place”, apresentando, igualmente, conteúdo que permite estabelecer relações entre essas situações e contextualizar as circunstâncias da sua ocorrência. A ordenação das situações torna-se, consequentemente, de elevada importância para a compreensão da complexidade temporal de um texto (cf. ISO, 2012).

Espera-se, por isso, que o *lead* relate a situação que motiva a notícia, mas, também, outras que procuram responder às questões referidas, expondo relações entre elas e contribuindo, do mesmo modo, para a compreensão da estrutura temporal do género.

É relevante, igualmente, mencionar o que Silva (2002: 184) argumenta na sua análise sobre os valores temporais nas sequências narrativas: “narrar consiste em estruturar verbalmente um conjunto de experiências, geralmente perspectivadas a partir de um intervalo de tempo posterior àquele em que ocorreram”. A narrativa é, por isso, e de acordo com Adam (1999 (2017)), uma sequência textual que encerra uma sucessão de situações e organiza o texto tendo em consideração as situações inicial e final e as fases narrativas que acontecem entre ambas.

3. Sobre tempo e aspeto

Dois dos conceitos fundamentais para a análise linguística das narrativas são as noções de tempo e de aspeto. Para este trabalho, consideraram-se, fundamentalmente, as questões que permitem esclarecer sobre a localização temporal das situações que ocorrem nas frases ou nas orações.

O tempo e o aspeto são duas categorias semânticas que se relacionam entre si, partilhando propriedades comuns, mas distintas a nível da sua classificação, ainda que interdependentes (Cunha, 2004). Oliveira (2003: 129) apresenta o primeiro como o localizador de “situações (eventos ou estados) expressas nas línguas em diferentes tipos de enunciados” e o segundo como a dimensão que “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”. O tempo, atómico, que permite ordenar, de forma linear, unidades temporais; o aspeto, subatómico, que perspetiva as situações a partir do seu interior (cf. Oliveira, 2003).

Reichenbach (1947) apresenta uma teoria bidimensional para descrever o tempo linguístico, expressa em três pontos fundamentais: (i) o tempo da fala (S), que corresponde ao momento de enunciação; (ii) o tempo da situação (E – *event*, em inglês), que diz respeito ao intervalo de tempo em que ocorre a situação; e (iii) o tempo de referência (R), que expressa o tempo tomado como referência e a partir do qual a situação é considerada.

Kamp & Reyle (1993), por seu turno, contribuem para a semântica das línguas naturais, particularmente no que diz respeito às questões de tempo e aspeto, à luz da *Discourse Representation Theory* (DRT). Dentro da semântica dinâmica, a DRT estabelece um ponto de perspetiva temporal (“temporal perspective time”), que permite localizar as situações, tentando desenvolver os pressupostos e limitações inerentes ao tempo de referência de Reichenbach (cf. Kamp et al., 2011: 199).

Saeed (2003) afirma que o tempo permite localizar uma situação relativamente a um ponto de referência numa linha temporal e diz respeito a um sistema dêitico cujo ponto de referência é,

² Silvano et al. (2021) fazem uso da palavra *events* considerando o contexto da ISO 24617 – *Language Resource Management – Semantic Annotation Framework (SemAF)* (ISO, 2012: 2), que declara o seguinte: “the term ‘event’ is used here with a very broad notion of event, which includes all kinds of actions, states, processes, etc. It is not to be confused with the more narrow notion of event as something that happens at a certain point in time (such as the clock striking 2, or waking up) or during a short period of time (such as laughing)”.

normalmente, o momento da enunciação. Nas palavras de Oliveira (2013: 510) “o tempo é concebido como um eixo linearmente ordenado” e “articula-se em três domínios, passado, presente e futuro, relativamente a um ponto ou a um intervalo tomado como referência, nomeadamente o momento em que o falante produz o enunciado”. É marcado na frase ou na oração por tempos verbais, adjuntos adverbiais de valor semântico temporal e verbos auxiliares ou semiauxiliares que operam no domínio do tempo.

O tipo de relações temporais que podem ser estabelecidas entre as situações representadas na frase levam em consideração a localização temporal da situação descrita, que se estabelece em função do ponto de referência temporal e que permite determinar relações de anterioridade ou posterioridade, quando está implícita uma relação sequencial, ou de simultaneidade, sempre que se verifica sobreposição temporal (cf. Oliveira, 2013).

4. Anotação da informação temporal

Ao nível da anotação da informação temporal, importa fazer referência à ISO 24617 – *Language Resource Management – Semantic Annotation Framework (SemAF)* (ISO, 2012), norma internacional criada para anotação semântica. Esta norma fornece diretrizes e especificações para a representação temporal estruturada e sistemática de dados linguísticos, promovendo a interoperabilidade entre diferentes sistemas de processamento de linguagem e esquemas de representação temporal. A parte um – *Language Resource Management – Semantic Annotation Framework (SemAF) – Part 1: Time and Events (SemAF-Time, ISO-TimeML)* (ISO 24617-1:2012) – está diretamente relacionada com o tempo e as situações (ou os *eventos*, a terminologia abrangente empregada pela norma) e foi usada como base para o desenho do esquema de anotação que conduziu o presente trabalho

A norma ISO 24617-1:2012 foi criada com base no sistema *TimeML* (e.o. Bunt & Pustejovsky, 2010; Pustejovsky et al., 2003a), uma linguagem de marcação desenvolvida para a anotação temporal de textos, que possibilita a identificação e classificação de informações temporais.

As diretrizes do *TimeML* foram aplicadas para a anotação de um *corpus* amplamente utilizado, o *TimeBank* (Pustejovsky et al., 2003b), cujo conteúdo foi utilizado para a formulação da norma ISO 24617-1:2012 (ISO, 2012). A versão portuguesa, apresentada por Costa & Branco (2012), foi apelidada de *TimeBankPT* e diz respeito ao primeiro *corpus* em português com anotações referentes, não apenas às expressões temporais, mas, também, às situações e às relações temporais estabelecidas entre elas, algo que, à data da publicação do estudo, constituiu uma novidade. Para criarem o *TimeBankPT*, Costa & Branco (2012) adaptaram os dados anotados com recurso ao *TimeML* e usados para o primeiro exercício de categorização de relações temporais em inglês, o *TempEval* (Verhagen et al., 2007).

5. O estudo

Neste ponto, começamos por descrever o *corpus* utilizado, fazendo menção ao *dataset* a partir do qual o mesmo foi recolhido (ponto 5.1). Seguidamente, passamos a apresentar o esquema de anotação adotado, partilhando partes do manual de anotação orientativo (ponto 5.2) e terminamos com a explicação detalhada do processo de anotação dos dados em análise para o presente estudo (ponto 5.3).

5.1. Descrição do corpus

Os dados recolhidos e apresentados neste trabalho correspondem a um *corpus* constituído por trinta notícias, seleccionadas aleatoriamente a partir de um *dataset* chamado *Text2Story Lusa* (Nunes et al., 2023).

Este *dataset*, desenvolvido no contexto do projeto *Text2Story: Extracting journalistic narratives from text and representing them in a narrative modeling language* (Campos et al., 2020), é formado por 358 notícias da agência Lusa, escritas em português europeu e publicadas entre outubro e dezembro de 2020, no site da agência³. O projeto *Text2Story*⁴ pretende extrair e representar as narrativas de notícias escritas em português europeu para facilitar a produção de *media* visual.

As notícias que integram o *dataset* são de natureza narrativa, com não mais de duzentas palavras cada, sendo que as suas temáticas variam entre o crime (na sua grande maioria), a política, o desporto e a sociedade local. O enfoque da análise contemplou, especificamente, a parte *lead* das notícias seleccionadas.

5.2. Esquema de anotação

O esquema de anotação utilizado para este estudo seguiu as orientações do manual de anotação orientativo do projeto *Text2Story*, que diz respeito a uma combinação e harmonização de cinco partes da norma ISO 24617 – *Language Resource Management – Semantic Annotation Framework (SemAF)* (apresentada anteriormente no ponto 4), proposta por Silvano et al. (2021). A parte da norma usada para este estudo constitui a ISO 24617-1:2012 (*SemAF-Time, ISO-TimeML*) (ISO, 2012) e refere-se ao tempo e aos eventos⁵.

O manual de anotação *Text2Story* introduz graficamente o esquema de anotação semântica multicamadas (ver figura 1), bem como, sucintamente, as fases do processo de anotação (apresentadas em detalhe no ponto 5.3.).

³ Disponível na *internet* em: <https://www.lusa.pt>.

⁴ Disponível na *internet* em: <https://storysense.inesctec.pt>.

⁵ Como mencionado anteriormente, *evento* constitui a terminologia usada na norma ISO e, do mesmo modo, no manual de anotação *Text2Story*, cuja definição corresponde a “eventualidade ou situação, de natureza dinâmica ou não dinâmica, que é temporalmente relevante”.

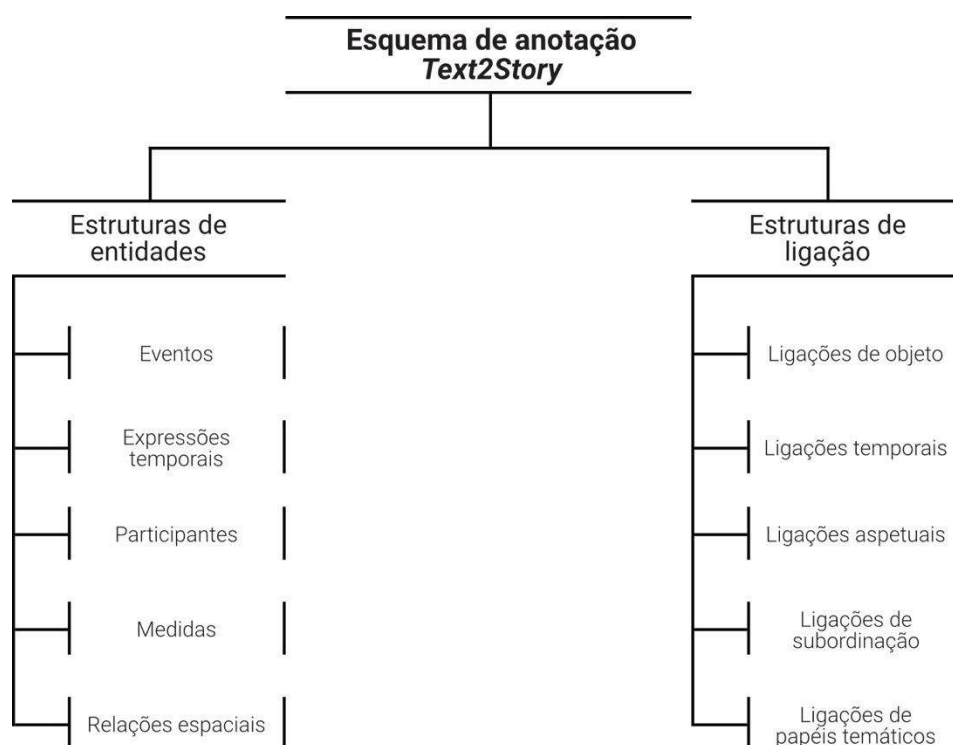


Figura 1: Esquema de anotação semântica multicamadas *Text2Story* (fonte: Silvano et al., 2023)

A secção correspondente às diretrizes de anotação divide-se em duas grandes áreas: (i) as estruturas de entidades, que se subdividem em eventos, expressões temporais, participantes, medidas e relações espaciais; e (ii) as estruturas de ligação, que se fragmentam em ligações de objeto, temporais, aspetuais, de subordinação e de papéis temáticos. É, também, nesta secção que se estabelece o que é “marcável” (*markable* – passível de ser anotado), determinando-se, igualmente, o valor dos atributos para o processo de anotação.

Para o presente estudo, foram utilizadas as etiquetas e os valores das estruturas de entidades e de ligação apresentados na figura 2 (ver abaixo). A saber: ao nível das estruturas de entidades, concretamente, em relação às situações, recorreu-se às etiquetas *POS* (classe gramatical), *tense* (tempo verbal) e *aspect* (valor aspetual), e respetivos valores; por outro lado, no que respeita às estruturas de ligação, utilizou-se a etiqueta *temporal links* e os valores correspondentes (BEFORE, AFTER, INCLUDES, IS_INCLUDED, SIMULTANEOUS, DURING, IDENTITY, BEGINS, ENDS, BEGUN_BY e ENDED_BY).

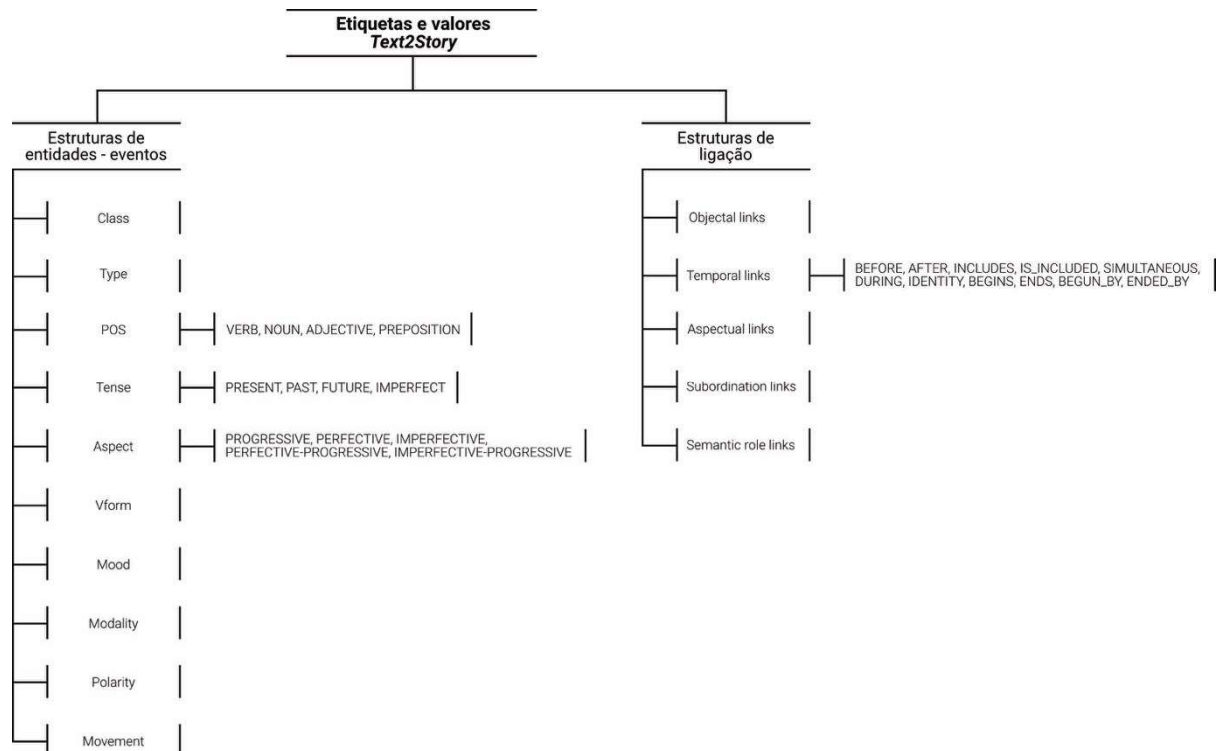


Figura 2: Etiquetas e valores das estruturas de entidades – eventos e das estruturas de ligação – ligações temporais (fonte: Silvano et al., 2023)

A tabela 1 apresenta as definições, tanto para as estruturas de entidades – eventos como para as estruturas de ligação – ligações temporais das etiquetas e dos valores a elas associados, dos atributos alvo da anotação. Para efeitos de facilitação da leitura, as definições que constam na tabela 1 serão retomadas em nota de rodapé, ao longo da parte da discussão.

Tabela 1

Definições dos valores dos atributos das estruturas de entidades – eventos e das estruturas de ligação – ligações temporais

Estruturas de entidades – eventos		
Atributos	parts of speech (POS)	VERB NOUN ADJECTIVE PREPOSITION
	tense	PRESENT – localiza a situação no momento presente, num futuro próximo ou num intervalo de tempo que começa no passado e se mantém no presente. Também pode representar situações habituais. PAST – localiza as situações num tempo passado. FUTURE – localiza as situações no futuro. IMPERFECT – localiza situações contínuas no passado.
	aspect	PROGRESSIVE – situação contínua. PERFECTIVE – situação terminada. IMPERFECTIVE – situação contínua, habitual ou iterativa. PERFECTIVE-PROGRESSIVE – combinação dos significados do progressivo e do perfeitivo. IMPERFECTIVE-PROGRESSIVE – combinação dos significados do progressivo e do imperfetivo.
Estruturas de ligação – ligações temporais		
Atributos	BEFORE – situação que ocorre antes de outra. AFTER – situação que ocorre depois de outra. INCLUDES – situação que inclui outra.	

IS_INCLUDED – situação que está incluída noutra.
 SIMULTANEOUS – situação que ocorre ao mesmo tempo que outra.
 DURING – situação que ocorre durante outra.
 IDENTITY – duas situações que correspondem à mesma situação.
 BEGINS – situação que começa outra.
 ENDS – situação que termina outra.
 BEGUN_BY – situação que é começada por outra.
 ENDED_BY – situação que é terminada por outra.

5.3. O processo de anotação

Para este estudo, a anotação das relações temporais entre as situações presentes no *lead* de notícias de agência tornou-se tarefa central. Para tal, e após definir-se o esquema de anotação a utilizar, o enfoque da análise contemplou o conteúdo do *lead* das notícias selecionadas.

A ferramenta de anotação e visualização de texto utilizada foi o BRAT (brat rapid annotation tool). Esta ferramenta foi desenvolvida e apresentada por Stenetorp et al. (2012) para suporte de anotação de dados em linguística computacional e tarefas de processamento da linguagem natural. O BRAT fornece uma interface web intuitiva que possibilita aos utilizadores rotular e destacar partes de texto de forma interativa (cf. Stenetorp et al., 2012).

O processo de anotação na ferramenta BRAT desenvolveu-se do seguinte modo: (i) numa primeira fase, efetuou-se um treino de anotação (com enfoque no procedimento descrito nas alíneas do ponto (ii), adiante), para aumento da competência de uso, tanto da ferramenta como do método de anotação, com recurso a dez notícias; (ii) ao longo da segunda fase, e após seleção aleatória das notícias, procedeu-se a:

- a) a identificação do *lead* de cada notícia;
- b) a identificação e anotação de todas as situações do *lead* de cada notícia;
- c) a seleção dos atributos de cada situação, de acordo com as etiquetas e valores das estruturas de entidades – eventos selecionados para este trabalho (cf. 5.2);
- d) o estabelecimento das relações temporais entre as situações de cada *lead*, da última para a imediatamente anterior, numa ordem linear, de acordo com as etiquetas e valores das estruturas de ligação – ligações temporais (cf. 5.2).

O *corpus* foi anotado considerando a classe gramatical, o tempo verbal e o valor aspetual das situações – *POS*, *tense* e *aspect* (cf. figura 2) e as relações temporais estabelecidas entre cada uma das situações identificadas e anotadas. Para este estudo, não foram consideradas as situações que representam relatos, como *anunciar*, *comunicar*, *informar*, *dizer*, *noticiar*, etc., nem as expressões temporais, como *hoje*, *o início da noite de terça-feira*, *a madrugada de hoje*, *o passado dia 13 de setembro*, etc.

Ao longo das duas primeiras fases do processo, a anotação realizada foi orientada e validada por um anotador mais experiente neste esquema de anotação e foram corrigidos alguns problemas.

A anotação obedeceu a critérios específicos elencados no manual do *Text2Story*. Os critérios mais relevantes para a elaboração deste trabalho são apresentados na lista abaixo:

- i. são apenas anotados os verbos principais (os verbos auxiliares não são considerados);
- ii. nos casos de construções predicativas, somente o adjetivo/nome/complemento preposicional é anotado (os verbos copulativos não são considerados), atribuindo-se ao adjetivo/nome/preposição os valores de *tense* e *aspect* do verbo copulativo;
- iii. se o verbo estiver incluído numa expressão (semi)lexicalizada (e.g. *será presente a tribunal*) marcar essa expressão como o predador;
- iv. a anotação dos nomes é feita sem especificadores, complementos e modificadores;

- v. no caso de nomes que representam situações e que são seguidos de modificadores adjetivais ou de um adjetivo formando uma expressão (semi)lexicalizada, marca-se a expressão completa (e.g. *fiscalização rodoviária*, *prisão efetiva*);
- vi. a anotação dos adjetivos é feita sem complementos e modificadores.
- vii. todas as situações retomadas posteriormente no texto serão anotadas, estabelecendo-se entre elas uma relação de identidade. Nos casos do pronome *que*, que retoma uma situação anterior (cf. (10)), é-lhe atribuído os valores do nome ao qual se liga.

Chegou-se, do mesmo modo, a consensos de análise e estabeleceram-se reajustes à anotação, como os exemplos apresentados seguidamente. Assim: (i) quando perante casos do tipo “crimes de X”, considerou-se “crimes” e “de X” como uma única situação (cf. (1)), por se tratar de um tipo de crime e não de duas situações distintas, por oposição a existirem duas situações diferentes ligadas pela relação de identidade; (ii) a relação estabelecida entre “prática” e “suspeito” foi de simultaneidade, uma vez que as duas situações ocorrem ao mesmo tempo (cf. (2)), e a relação entre “detido” e “suspeito” ou “suspeito” e “detido” foi de inclusão, pois o estado de ser “suspeito” inclui a culminação referente à detenção (cf. (3)); (iii) em contextos em que se verificou uma listagem de situações sem qualquer indicação da sua ordenação, optou-se por não se estabelecer relação temporal entre elas (cf. (4)). Por último, é de referir-se o excerto da notícia em (5), que exemplifica um contexto em que se verifica não estabelecimento deliberado de relação temporal entre duas situações. Neste caso, não há relação temporal entre [tentado] e [matar] mas, de facto, uma ligação de subordinação com o atributo INTENSIONAL⁶, que está fora do escopo deste trabalho.

As situações analisadas e expressas nos exemplos que serão dados ao longo das próximas páginas, estão assinaladas no texto das alíneas através do sublinhado. Por outro lado, as situações retiradas dos exemplos das alíneas e mencionadas no corpo de texto, são apresentadas entre parêntesis retos [].

- (1) Um homem de 71 anos foi detido e está em prisão preventiva, por fortes indícios da prática de crimes de abuso sexual de crianças (...). (lusa_60)
- (2) Um homem de 37 anos de idade, suspeito da prática de (...). (lusa_59)
[prática] – SIMULTANEOUS – [suspeito]
- (3) Um homem de 52 anos foi detido na ilha cabo-verdiana de São Vicente suspeito da autoria (...). (lusa_61)
[suspeito] – INCLUDES – [detido]
- (4) (...) por suspeita de burla qualificada e falsidade informática, praticados de forma continuada (...). (lusa_73)
- (5) (...) suspeito de ter tentado matar uma mulher (...). (lusa_62)

Tipicamente, o processo de anotação é feito da última situação para a imediatamente anterior, numa ordem linear, como ilustrado no manual de anotação. Contudo, casos houve em que a ligação mais relevante não seguiu esta ordem, como o exemplifica (6). De facto, a relação temporal é estabelecida entre [suspeita] e [deteve] ao invés de entre [suspeita] e [em flagrante delito], por existir entre os dois primeiros uma relação causal mais forte (cf. Setzer, 2001).

⁶ No manual de anotação *Text2Story*, o atributo INTENSIONAL está associado às estruturas de ligação – ligações de subordinação, cuja definição denota uma situação que introduz uma referência a mundos possíveis, principalmente I_ACTION e I_STATE (estruturas de entidades – eventos). (i) O João prometeu à Maria comprar uma casa nova.

- (6) A Polícia Judiciária (PJ) deteve em flagrante delito, nos arredores de Aveiro, dois indivíduos, de 19 e 36 anos, por suspeita de (...). (lusa_73)

Por via de regra, considerou-se, apenas, a leitura do *lead* para análise dos dados. No entanto, é relevante mencionar o caso da notícia expressa em (7), que expôs a necessidade de recorrer-se à leitura do corpo da notícia para a codificação da informação temporal, demonstrando-se, dessa forma, nem sempre ser possível processar toda a informação relevante apenas pela leitura do título e do *lead*. Segundo van Dijk (1983), as leituras do título e do *lead* serão suficientes para se processar a informação mais importante constante na notícia. No entanto, nesta situação, a leitura da informação contida no *lead* não permitiu, por si só, inferir que o momento da detenção foi anterior ao momento da declaração do homicida, como é possível observar-se em (7a); é a informação explicitada no corpo da notícia que clarifica a relação de posterioridade de [admitiu] face a [deteve], em (7b), contrapondo o que van Dijk defende.

- (7) a. A Polícia angolana deteve, na província da Huíla, um homem que admitiu ter matado a sua segunda mulher (...). (lusa_64, excerto do *lead*)
- b. (...) O suspeito, de 55 anos, foi detido apenas na quarta-feira. Fernando Tongo, citado pela agência noticiosa angolana Angop, explicou que, durante o interrogatório, o homem admitiu que matou a mulher (...). (lusa_64, excerto do corpo da notícia).

Decorridas todas as etapas elencadas acima, a anotação foi compilada e analisada e os seus resultados são apresentados no ponto que se segue.

6. Discussão e resultados

Nos *leads* das trinta notícias anotadas e analisadas, registou-se a ocorrência de 133 situações e de 104 relações temporais (representadas como TLINK), com uma média de 4,43 situações e de 3,46 TLINKs por *lead*, resultados que podem ser confirmados na Tabela 2. No que concerne ao número de palavras por *lead*, verifica-se uma média de 33,73, obedecendo-se ao valor orientado pela Lusa no seu livro de estilo (Lusa, 2019), de 35 palavras por *lead*. O valor máximo de palavras por *lead* foi de quarenta (ver lusa_57) e o mínimo de 25 (ver lusa_339, lusa_340, lusa_344). Face aos valores máximos e mínimos de situações e relações temporais estabelecidas, é possível observar: (i) um valor máximo de oito situações e de sete TLINKs, por *lead* (cf. (8)); e (ii) um valor mínimo de uma situação e zero TLINKs, por *lead* (cf. (9)), cuja circunstância constitui caso único em toda a amostra analisada.

- (8) O Presidente da República promulgou hoje o decreto da Assembleia da República que determina o uso obrigatório de máscara na rua, por um período de 70 dias, sempre que não seja possível cumprir o distanciamento físico recomendado. (lusa_337, oito situações, sete TLINKs)
- (9) A Unidade de Controlo Costeiro (UCC) da GNR apreendeu quase três toneladas de berbigão imaturo, em Vagos, no distrito de Aveiro (...). (lusa_34, uma situação, zero TLINKs)

Tabela 2**Dados estatísticos dos *leads* analisados, das suas situações e das relações temporais estabelecidas**

	<i>Lead</i>	Situações	<i>TLINKs</i>
Total	30	133	104
Média	33,73 (palavras/ <i>lead</i>)	4,43 (situações/ <i>lead</i>)	3,46 (<i>TLINKs</i> / <i>lead</i>)
Valores máximos	40 (palavras/ <i>lead</i>)	8 (situações/ <i>lead</i>)	7 (<i>TLINKs</i> / <i>lead</i>)
Valores mínimos	25 (palavras/ <i>lead</i>)	1 (situações/ <i>lead</i>)	0 (<i>TLINKs</i> / <i>lead</i>)

Considerando os dados expostos na Tabela 3, pode, igualmente, afirmar-se que, das 133 situações analisadas, 57 são nomes (representados pelo valor NOUN) (cf. (10)), 55 são verbos (representados pelo valor VERB) (cf. (10), (12) e (13)), 12 são adjetivos (representados pelo valor ADJECTIVE) e sete são preposições (representados pelo valor PREPOSITION) (cf. (11)). Verificaram-se, ainda, duas ocorrências (registadas na Tabela 3 como “Outros”) cuja anotação, face ao atributo *parts of speech* não é, à data da publicação deste trabalho, possível de representar na ferramenta BRAT, embora seja relevante mencioná-las aqui. As duas ocorrências correspondem ao sintagma nominal com função de predicativo do sujeito [outro na forma tentada] (ver lusa_61) e ao pronome [que] (cf. (10)).

Tabela 3**Total de ocorrência dos atributos das situações e das relações temporais analisadas**

Atributos			Total de ocorrências
Situações	parts of speech (POS)	VERB	55
		NOUN	57
		ADJECTIVE	12
		PREPOSITION	7
		(Outros)	2
	tense	PRESENT	5
		PAST	34
		FUTURE	0
		IMPERFECT	1
	aspect	PROGRESSIVE	0
		PERFECTIVE	42
		IMPERFECTIVE	1
		PERFECTIVE-PROGRESSIVE	0
		IMPERFECTIVE-PROGRESSIVE	0
	<i>TLINKs</i>		
	BEFORE	22	
	AFTER	10	
	INCLUDES	22	
	IS_INCLUDED	11	
	DURING	0	
	SIMULTANEOUS	15	
	IDENTITY	22	
	BEGINS	0	
	ENDS	0	
	BEGUN_BY	2	
	ENDED_BY	0	

A localização verbal mais frequente é o passado (representado pelo valor PAST⁷) (cf. (10)), presente em 34 situações. Anotaram-se, ainda, cinco situações com localização temporal no presente (representados pelo valor PRESENT⁸) (cf. (12) e (13)) e apenas um com *tense* no imperfeito

⁷ PAST – localiza as situações num tempo passado.

⁸ PRESENT – localiza a situação no momento presente, num futuro próximo ou num intervalo de tempo que começa no passado e se mantém no presente. Também pode representar situações habituais.

(representado pelo valor IMPERFECT⁹) (cf. (14)). As ocorrências relacionadas com PAST caracterizam-se por 33 verbos/sintagmas verbais e por um sintagma preposicional (ver lusa_57). No que diz respeito a PRESENT, as frequências de situações são representadas por dois sintagmas preposicionais (ver lusa_56, 60), dois verbos (ver lusa_337, lusa_344) e um adjetivo (ver lusa_337). A única ocorrência IMPERFECT refere-se a [coabitava] (cf. (13)), verbo *coabitar* no pretérito imperfeito do indicativo, na terceira pessoa do singular.

Em termos aspetuais, o perfeito (representado pelo valor PERFECTIVE¹⁰) é a ocorrência mais comum, registando-se em 42 situações (cf. (11)), em comparação com apenas uma ocorrência no imperfeito (representado como IMPERFECTIVE¹¹) (cf. (14)).

- (10) O chinês Quanhai Li foi hoje eleito para a presidência da Federação Internacional de Vela, em assembleia-geral realizada online, num escrutínio em que venceu o anterior presidente (...). (lusa_338)
[eleito] e [venceu]: POS: VERB, tense: PAST, aspect: PERFECTIVE
[realizada]: POS: VERB
[assembleia-geral] e [escrutínio]: POS: NOUN
[que]: (pronome)
- (11) Um homem de 71 anos foi detido e está em prisão preventiva, por fortes indícios da prática de crimes de abuso sexual (...). (lusa_60)
[em prisão preventiva]: POS: PREPOSITION
- (12) O Presidente da República promulgou hoje o decreto da Assembleia da República que determina o uso obrigatório de máscara (...). (lusa_337)
[determina]: POS: VERB, tense: PRESENT
- (13) (...) um homem de 34 anos “fortemente” indiciado pelo abuso sexual de uma menina de 12 anos, uma familiar com quem coabitava (...). (lusa_68)
[coabitava]: POS: VERB, tense: IMPERFECT, aspect: IMPERFECTIVE)

Os resultados obtidos mostram que a maioria dos tempos verbais das situações do *lead* corresponde ao pretérito perfeito do indicativo, corroborando os resultados apresentados por Tavares (1997) no seu estudo sobre o verbo no texto jornalístico. Segundo Cunha (2021: 112), o pretérito perfeito do indicativo constitui um “tempo do passado que localiza uma dada eventualidade terminada num intervalo anterior ao momento da enunciação”, sendo, por isso, considerado um tempo terminativo. Por seu turno, a ocorrência do pretérito imperfeito do indicativo, em (14), diz respeito a um estado, que se sobrepõe a um ponto de perspetiva temporal anterior ao do momento da enunciação (Cunha, 2004, 2021), traduzida no tempo verbal [coabitava].

De assinalar, também, o facto de, de entre todos os verbos que representam situações, o verbo *deter* ser o que se repete mais frequentemente (18 ocorrências), que surge tanto no pretérito perfeito do indicativo, na terceira pessoa do singular “deteve”, como no particípio passado “detido”. A natureza relacionada com a temática da criminalidade, e que caracteriza a maior parte das notícias que fazem parte do *corpus* deste trabalho, poderá justificar estas ocorrências. Verbos como *cometer*, *esfaquear*, *abusar*, *apreender* (cf. (14), (15), (16) e (17), respetivamente), entre outros, para além de apresentarem uma conotação muitas vezes atribuída ao crime, evidenciam, igualmente, o facto de as notícias lidarem com factos, informações precisas e objetividade (cf. Tavares, 1997).

⁹ IMPERFECT – localiza situações contínuas no passado.

¹⁰ PERFECTIVE – situação terminada.

¹¹ IMPERFECTIVE – situação contínua, habitual ou iterativa.

- (14) (...) 200 crimes de violação e um crime de abuso sexual de crianças, cometidos maioritariamente sobre as suas quatro filhas (...). (lusa_57)
- (15) (...) homem de 65 anos suspeito de ter esfaqueado outro (...). (lusa_71)
- (16) (...) ambos suspeitos de abusarem sexualmente de uma menor (...). (lusa_65)
- (17) A Unidade de Controlo Costeiro (UCC) da GNR apreendeu quase três toneladas (...). (lusa_34)

O facto de as notícias comunicarem situações localizadas num passado recente faria prever que grande parte dos tempos verbais correspondesse ao pretérito perfeito do indicativo e que fosse maioritariamente perfeitiva quanto ao aspeto, o que se verifica. Do mesmo modo se compreende que, sendo o *corpus* usado para este trabalho constituído por notícias de natureza narrativa, o tempo verbal predominante seja o pretérito. A propriedade canónica de relação de anterioridade ao momento da enunciação prevalece clara, em concordância com o que Silva (2002) redige (cf. ponto 2).

Do ponto de vista de **relações temporais**, e de acordo com o que Bell (1991) declara, não é esperada no *lead* uma ordem cronológica de situações. Considerando os pressupostos de Setzer (2001), pode afirmar-se, de igual forma, que as relações temporais podem ocorrer entre um evento e uma expressão temporal ou entre dois ou mais eventos. Considerando os resultados da Tabela 3, as relações temporais mais frequentes são as que dizem respeito a relações de anterioridade e de inclusão (representadas pelos valores BEFORE¹² e INCLUDES¹³, respetivamente) (cf. (18)), bem como a relações de identidade (representadas pelo valor IDENTITY¹⁴) (cf. (19)), presentes, cada uma, em 22 das 104 relações temporais que se estabelecem entre as situações. Seguem-se as relações de simultaneidade (representadas pelo valor SIMULTANEOUS¹⁵), que surgem 15 vezes nos *leads* analisados (cf. (19)). As relações de sucessividade temporal (representadas pelo valor AFTER¹⁶) (cf. (19) e (20)) verificam-se em, somente, dez das 104 relações temporais analisadas, confirmando a proposta de Bell (2006) de não sucessividade temporal. Registam-se, ainda, 11 relações representadas pelo valor IS_INCLUDED¹⁷ (cf. relação INCLUDES e (21)) e duas representadas pelo valor BEGUN_BY¹⁸ (cf. (22)).

- (18) A GNR deteve três homens por suspeitas de sequestro à mão armada de um homem de 42 anos (...). (lusa_70)
[sequestro à mão armada] – BEFORE – [suspeitas]
[suspeitas] – INCLUDES – [deteve]
- (19) O Presidente da República promulgou hoje o decreto da Assembleia da República que determina o uso obrigatório de máscara na rua, por um período de 70 dias, sempre que não seja possível cumprir o distanciamento físico recomendado. (lusa_337)
[recomendado] – SIMULTANEOUS – [distanciamento físico]
[distanciamento físico] – IDENTITY – [cumprir]
[cumprir] – SIMULTANEOUS – [possível]
[possível] – SIMULTANEOUS – [uso]
[obrigatório] – IS_INCLUDED – [uso]
[uso] – IDENTITY – [determina]

¹² BEFORE – situação que ocorre antes de outra.

¹³ INCLUDES – situação que inclui outra.

¹⁴ IDENTITY – situações que correspondem à mesma situação.

¹⁵ SIMULTANEOUS – situação que ocorre ao mesmo tempo que outra.

¹⁶ AFTER – situação que ocorre depois de outra.

¹⁷ IS_INCLUDED – situação que está incluída noutra.

¹⁸ BEGUN_BY – situação que é começada por outra.

- [determina] – AFTER – [promulgou]
- (20) O homem de 62 anos suspeito de matar a ex-companheira com uma arma de fogo em Grijó, concelho de Vila Nova de Gaia, na quarta-feira, foi detido pela Polícia Judiciária (PJ) (...). (lusa_69)
 [detido] – AFTER – [matar]
 [matar] – BEFORE – [suspeito]
- (21) (...) em estado muito grave foram hoje encontrados (...). (lusa_6)
 [encontrados] – IS_INCLUDED – [em estado muito grave]
- (22) (...) tomou hoje posse como Presidente da Bolívia, para os próximos cinco anos, sucedendo no cargo a Jeanine Áñez. (lusa_339)
 [sucedendo] – BEGUN_BY – [posse]

Os resultados encontrados permitiram, igualmente, perceber a variedade de relações temporais que podem estabelecer-se num mesmo *lead*. Exemplo disso é o *lead* parcialmente apresentado em (23), em que se verificam as relações de inclusão, entre [liberdade sexual] e [crimes] e entre [fortes indícios] e [em prisão preventiva], de identidade, entre [crimes] e [prática], de simultaneidade, entre [prática] e [fortes indícios], e de posterioridade, entre [em prisão preventiva] e [detido].

- (23) (...) detido e encontra-se já em prisão preventiva por fortes indícios da prática de dezenas de crimes contra a liberdade sexual (...). (lusa_56)
 [liberdade sexual] – INCLUDES – [crimes]
 [crimes] – IDENTITY – [prática]
 [prática] – SIMULTANEOUS – [fortes indícios]
 [fortes indícios] – IDENTITY – [em prisão preventiva]
 [em prisão preventiva] – AFTER – [detido]

A narrativa é uma sequência textual que se caracteriza por uma organização, que pode ser cronológica, das situações representadas (cf. Silva, 2002). Os resultados da análise deste trabalho permitiram concluir que tal fenómeno não se verifica nos *leads* das notícias, tal como já havia sido mencionado anteriormente, com a referência a Bell (1991, 2006). Na amostra analisada, as evidências são claras, como facilmente o ilustra o conteúdo da Tabela 3.

Ainda que o uso da etiqueta TLINK e dos seus valores tenha permitido, sem grande dificuldade e na sua grande maioria, especificar as relações que se estabeleceram entre as várias situações do *lead*, registaram-se, todavia, circunstâncias em que o processo de anotação necessitou de maior ponderação para se determinarem as relações implicadas. Os exemplos que se seguem são disso prova:

- i. **A escolha entre relação de anterioridade e relação de simultaneidade:** estabeleceu-se uma relação de anterioridade nas situações que diziam respeito a circunstâncias referentes a um crime/vários crimes, isolados ou não, mas facilmente circunscritos e de menor duração (cf. (24)). Por seu turno, em situações como a de (25), que diziam respeito a uma “prática”, i.e., uma situação continuada, de maior duração, a relação determinada foi a de simultaneidade.

- (24) (...) suspeito de mais de 200 crimes de violação e um crime de abuso sexual de crianças (...). (lusa_57)
 [crime de abuso sexual] – BEFORE – [suspeito]
 [crimes de violação] – BEFORE – [suspeito]
- (25) (...) suspeito da prática de crimes de coação sexual (...). (lusa_59)
 [prática] – SIMULTANEOUS – [suspeito]

- ii. **A escolha entre os valores INCLUDES, DURING e SUMULTANEOUS devido à sua proximidade semântica:** constatou-se uma dificuldade em circunscrever os limites referentes a situações de INCLUDES e DURING.

A norma, na qual o manual de anotação se baseia, também não é muito esclarecedora. Senão, veja-se: “One including the other (INCLUDES): as is the case between the temporal expression and the event in the following example: (82) John arrived in Boston last Thursday.” *versus* “One holds during the other (DURING): Similar to INCLUDES, but used to relate one event included within another event. (83) Mary sneezed while running.” (ISO, 2012: 57). Para o valor SUMULTANEOUS, a norma assume a seguinte definição: “Two events are judged simultaneous if they happen at the same time, or are temporally indistinguishable in context, i.e., they occur close enough so that further distinguishing their times makes no difference to the temporal interpretation of the text. This is also used for expressing the duration of an ongoing event, as in (79) Mary taught from 2 to 4.” (ibidem: 56).

De volta ao *corpus*, no caso (26), optou-se pelo estabelecimento da relação de inclusão entre [crimes de homicídio qualificado na forma tentada] e [festa familiar], considerando o facto de haver uma tentativa de homicídio que tentou ser concretizada por parte de um participante a outro, dentro do contexto de uma situação que estava, já, a decorrer (a [festa familiar]).

- (26) (...) deteve um homem de 34 anos suspeito de dois crimes de homicídio qualificado, na forma tentada, durante uma festa familiar (...). (lusa_67)
[festa familiar] – INCLUDES – [crimes de homicídio qualificado na forma tentada]

No ponto 5.3, foi mencionado o facto de nem sempre se verificar uma anotação das relações temporais feita da última situação para a imediatamente anterior, na ordem linear. Tal como Silva (2002: 184) declara “uma narrativa consiste basicamente numa sucessão de eventos, processo caracterizado por relações de causalidade”. E são estas que motivam a não linearidade relacional dos exemplos (27) e (28), o que constitui uma exceção face ao que o manual preconiza: em (27), a relação temporal é estabelecida entre [em prisão preventiva] e [detido]; em (28), entre [detido] e [suspeito].

- (27) (...) foi detido por fortes indícios da prática de mais de 700 crimes de abusos sexuais a uma criança atualmente com 9 anos, encontrando-se já em prisão preventiva (...). (lusa_58)
[em prisão preventiva] – AFTER – [detido]
- (28) (...) suspeito da prática de crimes de coação sexual, importunação sexual e tentativa de violação em agosto na área metropolitana de Lisboa, foi detido (...). (lusa_59)
[detido] – IS_INCLUDED – [suspeito]

7. Considerações finais

A análise temporal das notícias permite investigar a forma como as situações se organizam e estruturam ao longo do texto e elucida sobre o tipo de relações temporais que podem estabelecer-se entre essas situações.

Com este estudo, procurou-se estudar as relações temporais que se estabelecem entre as situações do *lead* de notícias de agência com carácter narrativo. Para tal, construiu-se um *corpus*, que foi anotado

considerando a classe gramatical, o tempo verbal e o valor aspetual das situações, bem como o estabelecimento das relações temporais.

Constatou-se que o *lead* das notícias se caracteriza por possuir situações representadas, predominantemente, por nomes e verbos, ainda que se tenham verificado algumas ocorrências de adjetivos, preposições e, ainda, um sintagma nominal com função de predicativo do sujeito e um pronome. Os verbos do *lead* encontram-se, maioritariamente, no pretérito perfeito do indicativo (valor PAST) e apresentam aspeto perfetivo (valor PERFECTIVE).

As relações temporais que se estabelecem entre as situações dos *leads* analisados disseram respeito a relações com os valores de BEFORE, INCLUDES, IDENTITY, SIMULTANEOUS, AFTER, IS_INCLUDED e BEGUN_BY, sendo as mais frequentes as relacionadas com anterioridade, inclusão, identidade e simultaneidade.

Apesar de as narrativas se caracterizarem por apresentar uma organização cronológica das suas situações, tal não se verifica na narrativa do *lead* das notícias: a baixa ocorrência de relações de posterioridade leva a afirmar-se que, de um modo geral, se não verifica sucessividade temporal no *lead*.

No *lead*, relata-se a situação principal juntamente com as situações que a contextualizam e justificam; a situação principal é, assim, retomada com frequência, de forma a responder-se às perguntas a que o *lead* se propõe responder, o que poderá justificar, igualmente, as relações de identidade verificadas.

Verificou-se, igualmente, a necessidade de se descrever em maior detalhe e contrastar as definições das relações temporais com os valores de INCLUDES, DURING e SIMULTANEOUS (por apresentarem semelhanças nas suas designações e, conseqüentemente, poderem levar a ambiguidades durante o processo de anotação), o que constituirá trabalho futuro. Do mesmo modo, pretendemos estender, futuramente, a análise ao corpo da notícia, analisando-se, por exemplo, se as situações do *lead* são retomadas no corpo da notícia, e de que forma, ou que tipo de informações adicionais são introduzidas e como são apresentadas. Em trabalho posterior, incluir-se-ão as expressões temporais bem como as situações de relato e as relações de ambas com as situações representadas nas notícias.

Ficam por explorar situações que criam contextos intensionais, como [tentar matar], que podem expressar uma vontade ou um desejo, sem necessariamente implicarem a sua concretização. Situações como as expressadas pelo sintagma preposicional [sem o seu consentimento], também representam circunstâncias que implicam uma não concretização. A reflexão futura sobre o tratamento deste tipo de situações no processo de anotação é pertinente, pois a sua consideração terá implicações, também, na representação visual da narrativa, um dos domínios para os quais se pretende contribuir.

Para além do contributo semântico do escopo do trabalho, os resultados encontrados podem, igualmente, contribuir para o campo das ciências de computação, uma vez que o tratamento da informação temporal é relevante para o desenvolvimento de métodos de extração de linhas temporais narrativas e para a criação de visualizações que permitam representar situações de um modo relacionado e causal.

Agradecimentos

Um agradecimento especial à Professora Doutora Purificação Silvano, docente na Faculdade de Letras da Universidade do Porto e minha orientadora, e à equipa do *Text2Story/StorySense*, pelo apoio constante na elaboração deste trabalho.

Referências

- Adam, J. M. (1999 (2017)). *Les textes: types et prototypes*. Armand Colin. Disponível em <https://www.cairn.info/les-textes-types-et-prototypes--9782200617288.htm>.
- Bell, A. (1991). *The language of news media*. Blackwell.
- Bell, A. (1995). News time. *Time and Society*, 4(3), 305–328.
- Bell, A. (2006). News language. In K. Brown (Ed.), *Encyclopedia of language & linguistics* (pp. 615–617). Elsevier. <https://doi.org/10.1016/B0-08-044854-2/00716-1>
- Bendler, J. R., Davenport, L. D., Drager, M. W. & Fedler, F. (2018). *Writing and reporting for the media* (12th ed.). Oxford University Press.
- Bunt, H. & Pustejovsky, J. (2010). Annotating temporal and event quantification. *Proceedings of the Fifth Joint ACL-ISO Workshop on Interoperable Semantic Annotation ISA-5*, 15–22. Hong Kong.
- Campos, R., Jorge, A. M., Jatowt, A., Bhatia, S., Pasquali, A., Cordeiro, J. P., Rocha, C., Mansouri, B. & Santana, B. (2020). *Report on the third international workshop on narrative extraction from texts* (Text2Story 2020).
- Carnaz, G., Antunes, M. & Nogueira, V. B. (2021). An annotated corpus of crime-related portuguese documents for NLP and machine learning processing. *Data*, 6(7). <https://doi.org/10.3390/data6070071>
- Costa, F. & Branco, A. (2012). TimeBankPT: A TimeML annotated corpus of portuguese. In N. Calzolari, K. Choukri, T. Declerck, M. Uğur Doğan, B. Maegaard, J. Mariani, A. Moreno, J. Odijk & S. Piperidis (Eds.), *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)* (pp. 3727–3734). Istanbul.
- Cunha, L. F. (2004). *Semântica das predicções estativas para uma caracterização aspetual dos estados* [Dissertação de doutoramento, Faculdade de Letras da Universidade do Porto]. Repositório Aberto da Universidade do Porto. <https://repositorio-aberto.up.pt/handle/10216/18379>
- Cunha, L. F. (2021). Tempos do passado e tempos do futuro em confronto: diferenças e paralelismos. *Revista da Associação Portuguesa de Linguística*, 8, 105–120.
- International Organization for Standardization (2012). *Language Resource Management – Semantic Annotation Framework (SemAF) – Part 1: Time and Events (SemAF-Time, ISO-TimeML)* (ISO 24617-1:2012). <https://www.iso.org/standard/37331.html>
- Kamp, H., van Genabith, J. & Reyle, U. (2011). Discourse representation theory. In D.M. Gabbay & F. Guenther (Eds.), *Handbook of philosophical logic* (Vol. 15, pp. 124–394). Springer.
- Kamp, H. & Reyle, U. (1993). *From discourse to logic – Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory* (Vol. 42). Springer.
- Lusa – Agência de Notícias de Portugal (2019). *Livro de estilo*. Retirado de <https://www.lusa.pt/Files/lusamaterial/PDFs/LivroEstilo.pdf>.
- 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>
- Público (2005). *Livro de estilo* (2.^a ed.). Retirado de <https://static.publico.pt/files/provadosfactos/livro-de-estilo.pdf>.
- Oliveira, F. (2003). Tempo e aspeto. In M. H. M. Mateus, A. M. Brito, I. Duarte, I. H. Faria, S. Frota, G. Matos, F. Oliveira, M. Vigário & A. Villalva (Eds.), *Gramática da língua portuguesa* (pp. 127–168). Editorial Caminho.
- Oliveira, F. 2013. Tempo verbal. In M. B. P. Raposo, M. F. B. Nascimento, M. A. C. Mota, L. Segura & A. Mendes (Orgs.), *Gramática do português* (pp. 509–553). Fundação Calouste Gulbenkian.
- Pustejovsky, J., Castaño, J., Ingria, R., Sauri, R., Gaizauskas, R., Setzer, A., Katz, G. & Radev, D. (2003a). TimeML: Robust specification of event and temporal expressions in text. *IWCS-5, Fifth International Workshop on Computational Semantics*, 28–34. Tilburg.
- Pustejovsky, J., Hanks, P., Sauri, R., See, A., Gaizauskas, R., Setzer, A., Radev, D., Sundheim, B., Day, D., Ferro, L. & Lazo, M. (2003b). The TIMEBANK corpus. *Proceedings of the Corpus Linguistics 2003 conference*, 647–656. Lancaster.
- Reichenbach, H. (1947). *Elements of symbolic logic*. Macmillan.
- Saeed, J. I. (2003). *Semantics* (2nd ed.). Blackwell Publishing.
- Setzer, A. (2001). *Temporal information in newswire articles: An annotation scheme and corpus study* [PhD Thesis, University of Sheffield].
- Silva, P. N. (2002). A expressão de valores temporais numa sequência narrativa. *Vidya*, 37, 179–195.

- Silvano, P., Amorim, E., Leal, A., Cantante, I., Silva, F., Jorge, A., Campos, R. & Nunes, S. (2023). Annotation and visualization of reporting events in textual narratives. In R. Campos, A. Jorge, A. Jatowt, S. Bhatia & M. Litvak (Eds.), *Proceedings of the Text2Story'23 Workshop* (pp. 47–62). Dublin.
- Silvano, P., Leal, A., Silva, F., Cantante, I., Oliveira, F. & Jorge, A. (2021). Developing a multilayer semantic annotation scheme based on ISO standards for the visualization of a newswire corpus. *Proceedings of the 17th Joint ACL – ISO Workshop on Interoperable Semantic Annotation*, 1–13. Groningen.
- Stenetorp, P., Pyysalo, S., Topić, G., Ohta, T., Ananiadou, S. & Tsujii, J. (2012). BRAT: a web-based tool for NLP-assisted text annotation. *Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics*, 102–107. Avignon: Association for Computational Linguistics. Retrieved from <https://www.aclweb.org/anthology/E12-2021.pdf>
- Tavares, M. A. (1997). O verbo no texto jornalístico: a notícia e a reportagem. *Working Papers em Linguística*, 1, 123–142.
- van Dijk, Teun A. (1983). Discourse analysis: its development and application to the structure of news. *Journal of Communication*, 33(2), 20–43.
- van Dijk, Teun A. (1985). Structures of news in the press. In T. A. van Dijk (Ed.), *Discourse and communication: new approaches to the analysis of mass media discourse and communication* (Vol. 10, pp. 69–93). de Gruyter.
- van Dijk, Teun A. (1988). *News as discourse*. Lawrence Erlbaum Associates.
- Verhagen, M., Gaizauskas, R., Schilder, F., Hepple, M. & Pustejovsky, J. (2007). SemEval-2007 task 15: TempEval temporal relation identification. *Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007)*, 75–80. Association for Computational Linguistics.

Anexo A – Corpus

<i>Designação da notícia</i>	<i>Título</i>	<i>Lead</i>
lusa_56	<i>Dirigente de associação de apoio a carenciados detido por suspeita de crimes sexuais</i>	Um dirigente de uma associação de apoio a pessoas carenciadas foi detido e encontra-se já em prisão preventiva por fortes indícios da prática de dezenas de crimes contra a liberdade sexual de seis vítimas, anunciou hoje a Polícia Judiciária.
lusa_57	<i>Prisão preventiva para suspeito de 200 crimes de violação sexual sobre as filhas</i>	Um homem de 55 anos ficou em prisão preventiva por ser suspeito de mais de 200 crimes de violação e um crime de abuso sexual de crianças, cometidos maioritariamente sobre as suas quatro filhas, anunciou hoje a Polícia Judiciária (PJ).
lusa_58	<i>Detido homem por mais de 700 crimes de abuso sexual a criança de 9 anos</i>	Um homem, de 49 anos, foi detido por fortes indícios da prática de mais de 700 crimes de abusos sexuais a uma criança atualmente com 9 anos, encontrando-se já em prisão preventiva, anunciou hoje a Polícia Judiciária (PJ).
lusa_59	<i>Detido homem por suspeita de crimes de natureza sexual em Lisboa</i>	Um homem de 37 anos de idade, suspeito da prática de crimes de coação sexual, importunação sexual e tentativa de violação em agosto na área metropolitana de Lisboa, foi detido pela Polícia Judiciária (PJ), comunicou hoje esta força policial.
lusa_6	<i>Um morto e um ferido grave encontrados em Peniche com ferimentos de arma de fogo</i>	Um homem morto e uma mulher em estado muito grave foram hoje encontrados numa caravana, em Peniche, no distrito de Leiria, ambos com ferimentos de arma de fogo, disse à Lusa fonte policial.
lusa_60	<i>Homem de 71 anos em prisão preventiva por crimes sexuais contra crianças e adolescentes</i>	Um homem de 71 anos foi detido e está em prisão preventiva, por fortes indícios da prática de crimes de abuso sexual de crianças, atos sexuais com adolescentes e importunação sexual de duas menores, anunciou hoje a Polícia Judiciária.
lusa_61	<i>Homem detido na ilha cabo-verdiana de São Vicente por crimes de homicídio</i>	Um homem de 52 anos foi detido na ilha cabo-verdiana de São Vicente suspeito da autoria de um crime de homicídio e de outro na forma tentada, divulgou hoje a Polícia Judiciária (PJ).
lusa_62	<i>PJ deteve suspeito de tentar matar mulher em Vila Real</i>	A Polícia Judiciária (PJ) deteve um homem de 24 anos, em Vila Real, suspeito de ter tentado matar uma mulher e de a atingir por “diversas vezes” e “violentamente” com uma faca, anunciou hoje aquela força policial.
lusa_63	<i>GNR deteve em Paredes dois homens por furto em estaleiro de construção civil</i>	A GNR de Paredes anunciou hoje ter detido, em flagrante delito, dois homens de 30 e 35 anos, por furto qualificado em estaleiro de proteção civil, naquele concelho do distrito do Porto.
lusa_64	<i>Polícia angolana deteve homem que matou a mulher por gastar 3,9 euros sem permissão</i>	A Polícia angolana deteve, na província da Huíla, um homem que admitiu ter matado a sua segunda mulher, por esta supostamente ter gasto 3.000 kwanzas (3,9 euros) sem o seu consentimento.
lusa_65	<i>PJ deteve mãe de menina de 13 anos e o ex-companheiro por abuso sexual da menor</i>	A Polícia Judiciária (PJ) de Setúbal deteve um homem de 60 anos e a ex-companheira, de 50, ambos suspeitos de abusarem sexualmente de uma menor, filha do elemento feminino do casal, foi hoje anunciado.
lusa_66	<i>PSP do Porto deteve três homens suspeitos de exploração ilícita de jogo</i>	Três homens foram detidos em Vila Nova de Gaia e na Maia no âmbito de operações policiais relacionadas com o combate à prática de jogo de fortuna ou azar, revelou hoje o Comando Metropolitano da PSP do Porto.
lusa_67	<i>Polícia Judiciária deteve de dupla tentativa de homicídio em Almada</i>	A Polícia Judiciária (PJ) deteve um homem de 34 anos suspeito de dois crimes de homicídio qualificado, na forma tentada, durante uma festa familiar no passado dia 13 de

		setembro, em Almada, no distrito de Setúbal, foi hoje anunciado.
lusa_68	<i>PJ deteve suspeito de abusar de menina de 12 anos em Ílhavo</i>	A Polícia Judiciária (PJ) de Aveiro deteve um homem de 34 anos “fortemente” indiciado pelo abuso sexual de uma menina de 12 anos, uma familiar com quem coabitava, informou hoje aquela polícia.
lusa_69	<i>Polícia Judiciária deteve suspeito de matar ex-mulher em Vila Nova de Gaia</i>	O homem de 62 anos suspeito de matar a ex-companheira com uma arma de fogo em Grijó, concelho de Vila Nova de Gaia, na quarta-feira, foi detido pela Polícia Judiciária (PJ), revelou hoje aquela força policial.
lusa_70	<i>GNR deteve três homens suspeitos de sequestro na Costa de Caparica</i>	A GNR deteve três homens por suspeitas de sequestro à mão armada de um homem de 42 anos, na madrugada de hoje na Costa de Caparica.
lusa_71	<i>PJ deteve suspeito de esfaquear homem em Braga</i>	A Polícia Judiciária (PJ) deteve um homem de 65 anos suspeito de ter esfaqueado outro homem em Braga, ao início da noite de terça-feira, anunciou hoje aquela força policial.
lusa_72	<i>GNR deteve nove pessoas no distrito de Aveiro por suspeita de tráfico de droga</i>	A GNR deteve em Oliveira de Azeméis e Albergaria-a-Velha, no distrito de Aveiro, nove pessoas, com idades entre os 24 e 57 anos, por suspeita de tráfico de droga, informou hoje aquela força militar.
lusa_73	<i>PJ deteve dupla em Aveiro por burla na Internet</i>	A Polícia Judiciária (PJ) deteve em flagrante delito, nos arredores de Aveiro, dois indivíduos, de 19 e 36 anos, por suspeita de burla qualificada e falsidade informática, praticados de forma continuada, informou hoje aquele órgão de polícia criminal.
lusa_74	<i>GNR deteve homem por posse de 160 doses de canábis em Seia</i>	A GNR deteve um homem de 28 anos, por alegado tráfico de estupefacientes, que estava na posse de 160 doses de canábis, na localidade de Loriga, no concelho de Seia, foi hoje anunciado.
lusa_337	<i>Presidente da República promulga uso obrigatório de máscara na rua por 70 dias</i>	O Presidente da República promulgou hoje o decreto da Assembleia da República que determina o uso obrigatório de máscara na rua, por um período de 70 dias, sempre que não seja possível cumprir o distanciamento físico recomendado.
lusa_338	<i>Chinês Quanhai Li eleito presidente da Federação internacional de vela</i>	O chinês Quanhai Li foi hoje eleito para a presidência da Federação Internacional de Vela, em assembleia-geral realizada online, num escrutínio em que venceu o anterior presidente, o dinamarquês Kim Andersen.
lusa_339	<i>Luis Arce tomou posse como Presidente da Bolívia</i>	O economista Luís Arce, de 57 anos, tomou hoje posse como Presidente da Bolívia, para os próximos cinco anos, sucedendo no cargo a Jeanine Áñez.
lusa_34	<i>GNR apreendeu cerca de três toneladas de berbigão imaturo em Vagos</i>	A Unidade de Controlo Costeiro (UCC) da GNR apreendeu quase três toneladas de berbigão imaturo, em Vagos, no distrito de Aveiro, informou hoje aquela força militar.
lusa_340	<i>Ex-Presidente francês Giscard d'Estaing foi hospitalizado no domingo</i>	O ex-Presidente francês Valéry Giscard d'Estaing, 94 anos, foi hospitalizado no domingo, noticiam hoje os ‘media’ franceses, que não especificam o motivo do seu internamento.
lusa_341	<i>Presidente da federação haitiana banido pela FIFA por abuso sexual</i>	O presidente da Federação Haitiana de Futebol, Yves Jean-Bart, foi banido para sempre de toda a atividade ligada ao futebol, por ter sido considerado culpado de abuso sexual sistemático de jogadoras, anunciou hoje a FIFA.
lusa_342	<i>Vice-presidente angolano em São Tomé para funeral do antigo presidente do parlamento</i>	O vice-presidente de Angola, Bornito de Sousa Diogo, chegou esta manhã a São Tomé e Príncipe para participar hoje nas cerimónias fúnebres do antigo presidente do parlamento são-tomense, Alcino Pinto, falecido na quinta-feira vítima de paragem cardiovascular.

lusa_343	<i>Morreu Manuel Nobre, ex-presidente da Câmara de Rio Maior e fundador das Carnes Nobre</i>	O antigo presidente da Câmara Municipal de Rio Maior, Manuel Sequeira Nobre morreu sexta-feira, aos 93 anos, tendo o município decretado três dias de luto municipal, disse hoje à Lusa o atual presidente, Filipe Dias.
lusa_344	<i>Presidente da República do Congo em Bissau na terça-feira para visita oficial</i>	O Presidente da República do Congo, Denis Sassou Nguesso, realiza terça e quarta-feira uma visita oficial à Guiné-Bissau, informou hoje, em comunicado, a Presidência guineense.
lusa_345	<i>Covid-19: Presidente francês infectado pelo novo coronavirus - Oficial</i>	O Presidente francês, Emmanuel Macron, testou positivo para a covid-19, após ter apresentado os “primeiros sintomas da doença”, anunciou a presidência, diagnóstico confirmado hoje, um dia depois de uma reunião com o primeiro-ministro português, António Costa.