

Uncovering Temporal Framing in the News

Tarek Mahmoud¹, Veronika Solopova^{2,3}, Premtim Sahitaj^{2,3}, Ariana Sahitaj^{2,3},
Max Upravitelev^{2,3}, Mervat Abassy¹, Hana Fatima Shaikh⁴, Neda Foroutan²,
Vera Schmitt^{2,3}, Preslav Nakov¹

¹MBZUAI, UAE ²Technische Universität Berlin, Germany

³German Research Center for Artificial Intelligence (DFKI), Germany

⁴University of Maryland, USA

{tarek.mahmoud, preslav.nakov}@mbzuai.ac.ae

Abstract

Temporal language does more than place events on a timeline. In news discourse, references to the past, present, and future can function as rhetorical devices that shape interpretation and persuasion. Here, we study *temporal framing*, defined as the persuasive use of time-related language to structure meaning rather than to report chronology. We propose a taxonomy of eight temporal frames grounded in prior work on temporality and framing, and we realize it through expert annotation of a multilingual news corpus. The resulting dataset includes 458 English and German news articles, with over 2K temporally framed sentences and approximately 3K temporal framing annotations identified from a corpus of more than 20K sentences. We analyze frame prevalence, co-occurrence patterns, and lexical cues, and evaluate temporal framing detection using supervised fine-tuning and zero-shot classification. Our experiments show that temporal framing is learnable at the sentence level, with supervised models substantially outperforming zero-shot approaches. We publicly release the corpus to support future research on temporal framing: <https://mbzuai-nlp.github.io/temporal-framing/>.

1 Introduction

All languages provide systematic means to express temporal relations, such as ordering, simultaneity, duration, and change over time, whether through tense, aspect, lexical items, or discourse-level mechanisms (Comrie, 1985; Evans and Levinson, 2009). Beyond merely locating events on a timeline, temporal expressions influence how speakers establish causality and how they invite audiences to reason about the past, the present, and the future. Linguistic choices related to time play a central role not only in describing reality, but also in framing it (Hartog and Brown, 2015).

In computational linguistics, however, time has traditionally been treated in a narrow instrumental manner, with focus on predicting verb tense or aspect, while recent work in NLP has concentrated on temporal expression extraction and temporal reasoning (Tan et al., 2023; Ding and Wang, 2025). Although these efforts have enabled significant progress in temporal reasoning, they typically conceptualize time as an objective property of events, and not a rhetorical resource. As a result, current NLP systems largely overlook the richer semantic and pragmatic functions that temporal language fulfills in discourse. While prior work has explored framing in NLP more broadly (Otmakhova et al., 2024; Mahmoud et al., 2025), the intersection of time and framing remains underexplored. This gap limits our ability to analyze persuasive rhetorical structures along the temporal dimension, particularly in domains such as political communication, where temporal justifications and appeals are known to be especially influential (Solopova et al., 2023). Concretely, temporal framing enables fine-grained analysis of how news and political discourse construct narratives through references to the past, present, and future, supporting downstream applications such as media profiling, stance and narrative analysis, and cross-lingual comparison of framing strategies across outlets and events.

In social sciences, time is widely recognized as a powerful framing device (Trope and Liberman, 2010). Temporal framing categories enable systematic study of phenomena such as appeals to nostalgia, constructions of urgency, or the use of historical anchoring to legitimize policy decisions across events, outlets, and time, in diverse communicative and sociopolitical contexts globally and historically. In political discourse studies, temporal references are explicitly analyzed as rhetorical strategies that frame interpretation and legitimize authority (Wodak, 2015).

Violence Has Been Normalized

During the misnamed and mostly preposterous debate between Kamala Harris and Donald Trump, a moderator fact-checked Trump’s claim that crime is up.

In contrast to Trump’s claim, moderator David Muir said that the FBI reports that crime is down, a claim that likely struck every viewer as obviously wrong.

Shoplifting was not a way of life before lockdowns. Nostalgia Temporal Anchoring Temporal Contrast Most cities were not demographic minefields of danger around every corner. Nostalgia Temporal Contrast There was no such thing as a drugstore with nearly all products behind locked Plexiglas. Nostalgia Temporal Contrast

We weren’t warned of spots in cities, even medium-sized ones, where carjacking was a real risk. Nostalgia Temporal Contrast

It is wildly obvious that high crime in the U.S. is endemic, with ever less respect for person and property. Continuity Primacy
 As for the FBI’s statistics, they’re worth about as much as most data coming from federal agencies these days. Temporal Contrast

(a) **Annotated excerpt illustrating temporal framing.** The complete example appears in Appendix F.

Temporal Frame	Short Definition	Example
Primacy	<i>Significance attributed to being first in time.</i>	“The first to discover the cure will lead the world.”
Recency	<i>Significance attributed to the most recent events.</i>	“Today’s figures matter most for understanding public opinion.”
Urgency	<i>Emphasis on limited time or imminent consequences or threats.</i>	“Scientists warn the next 48 hours are critical.”
Temporal Anchoring	<i>Framing the discussion through the lens of past events.</i>	“We live in a post-9/11 world where security comes first.”
Nostalgia	<i>Invocation of a cherished past.</i>	“We must act now to bring back the good old days.”
Temporal Contrast	<i>Juxtaposition of “then” versus “now.”</i>	“Once a booming hub, now a city in decline.”
Continuity	<i>Persistence across time.</i>	“The economy has been climbing steadily for a decade thanks to policy changes.”
Skeptical	<i>Casting doubt about the future.</i>	“The government’s plan may collapse under financial pressure.”

(b) **Brief overview of the temporal framing taxonomy with definitions and examples.** For detailed definitions, additional examples, and rhetorical functions, see Appendix B.

Figure 1: **An annotated excerpt and an overview of the temporal framing taxonomy.**

Temporal references can create urgency by compressing perceived time, convey stability by emphasizing continuity, evoke nostalgia of an idealized past, or anchor present arguments to emotionally charged historical moments. These patterns are routinely studied in domains such as political communication, crisis reporting, and health messaging, where temporal framing shapes interpretation and decision-making. The persuasive force of temporal references lies in their relational positioning within the discourse. Expressions such as *now* or *for decades* gain rhetorical meaning from implied alternative timeframes. Temporal framing operates at a rhetorical level, emerging from how temporal references are positioned within a sentence and the broader context. This makes temporal framing challenging to capture with purely extraction-based approaches.

Motivated by this gap, we present a computational study of temporal framing in news discourse (see Figure 1a for an example). Our contributions are fourfold:

1. We cast temporal framing as a sentence-level, multi-label task capturing the rhetorical use of temporal language beyond surface expressions.
2. We propose a taxonomy of eight temporal frames grounded in social science research.
3. We introduce an annotation scheme and release a multilingual news dataset of 458 articles annotated at the sentence level.
4. We provide computational baselines for detecting temporal framing using diverse language models in zero-shot and fine-tuned settings.

2 Related Work

Framing in NLP Framing has been widely studied in NLP. Early work has primarily focused on *document-level framing*, where each article is assigned one or more dominant frames (Card et al., 2015; Liu et al., 2019; Alonso del Barrio and Gatica-Perez, 2023). This line of work has enabled large-scale analysis of framing trends across outlets and over time spans, supporting comparative studies of media bias and agenda setting, but abstracts away from how framing is realized locally in language and interaction. More recent work has explored *entity-level*, *event-level*, and *narrative* framing, examining how framing emerges through the roles, attributes, and relationships assigned to entities and events (Stammbach et al., 2022; Otmakhova et al., 2024; Gehring and Grigoletto, 2023; Das et al., 2024; Mahmoud et al., 2025). These approaches capture finer-grained framing dynamics and provide insight into how meaning is constructed within texts, including the interplay between actors and events in discourse. While these approaches capture finer-grained framing dynamics, they do not explicitly model the rhetorical use of temporal language as a framing dimension.

Temporal Analysis in NLP A large body of research addresses temporal understanding through tasks such as temporal expression extraction, event ordering, timeline construction, and temporal reasoning (Chu et al., 2024; Islakoglu and Kalo, 2025). Benchmarks and models in this area aim to recover factual temporal structure, answering questions about when events occurred and how they relate chronologically. Although these methods advance computational representations of time, they largely treat temporal information as objective and descriptive rather than persuasive.

Recent work has also examined temporal structure in news at the corpus level. Although Cobo et al. (2025) use the term *temporal framing*, their work defines temporality as the evolution of frames across time rather than as a rhetorical device expressed in language. Likewise, Card et al. (2022) analyze how news framing varies over time. Similarly, Lamorte et al. (2024) introduce an interactive system for event-centered exploration of news archives, enabling users to browse events and frames along a timeline. While these approaches capture how frames and events evolve over time, they treat time as an organizing or analytical dimension rather than a rhetorical one.

Temporal Framing in Social Sciences Temporal framing has been extensively studied in psychology, communication, and political science. Foundational work in construal level theory shows that temporal distance systematically shapes cognition, with proximal events construed more concretely and distal events more abstractly, influencing evaluation, emotion, and decision-making (Trope and Liberman, 2010). In applied domains such as health communication, temporal framing of risks and outcomes affects perceived severity and behavioral intentions (Chandran and Menon, 2004; Kim and Kim, 2018; Orbell and Kyriakaki, 2008). In political discourse, temporal references to the past or future are used to structure interpretation, justify policy choices, and legitimize action (Wodak, 2015; Wagoner and Herbig, 2025). Despite this extensive work, these insights have not yet been systematically analyzed or modeled in NLP.

Although *temporal framing* is well studied in social sciences, NLP has largely overlooked time as a rhetorical device. Existing datasets and models fail to capture how authors use time beyond factual chronology. We address this gap by introducing a sentence-level temporal framing taxonomy and annotated dataset for news discourse.

3 Temporal Framing

Temporal framing refers to the rhetorical use of time-related elements to structure meaning and persuade audiences. Rather than merely situating events chronologically, it highlights, contrasts, or attributes significance to moments, durations, and trajectories to shape how issues are perceived.

To make these distinctions more precise, we introduce a taxonomy of temporal frames summarized in Figure 1b and described in Appendix B. The taxonomy identifies eight ways temporal reference is used rhetorically, ranging from appeals to primacy, recency, or urgency to frames invoking nostalgia, continuity, or skepticism about the future. Temporal references may be accompanied by markers such as *first*, *now*, *still*, *since*, *before*, *after*, *deadline*, *turning point*, *for years*, *pre-war*, *last month*, *next year*. The taxonomy is vital for understanding nuanced uses of time, enabling granular rhetorical analysis, and is grounded in social science research (Appendix D). It reflects theoretically grounded and empirically stable patterns observed during annotation rather than an exhaustive set of temporal expressions.

Temporal framing can be formalized as follows. Let \mathcal{F} denote the set of temporal frames defined by the taxonomy. Let $D = (s_1, \dots, s_n)$ be a document consisting of n sentences. The goal of temporal framing is to learn a function

$$f : (D, s_i) \rightarrow \mathcal{P}(\mathcal{F}),$$

where s_i is a sentence in the document and $\mathcal{P}(\mathcal{F})$ denotes the power set of temporal frames. The output $f(D, s_i)$ corresponds to the set of temporal frames whose persuasive force is expressed through time-related language in sentence s_i , possibly conditioned on its surrounding context in D .

Identifying temporal framing requires resolving two recurring challenges: (1) distinguishing *factual* chronological reports from *opinionated* temporal framings, and (2) focusing on the *temporal* rather than the *general* sense of a frame.

Fact versus opinionated use. A central distinction in our analysis is between factual chronological reporting and opinionated temporal framing. Factual reports are verifiable statements that merely *describe* events or states, such as dated statistics or chronological markers, without seeking to persuade. Opinionated framings, by contrast, are rhetorical by nature: they *utilize temporal elements* to persuade readers. The same sentence can shift from a neutral description to a persuasive framing depending on whether the temporal element is presented as meaningful rather than incidental. For example, “Inflation rose by 2% in 2024” reports a temporal fact, whereas “Inflation has been rising steadily for years, signaling policy failure” uses temporal reference to support a persuasive claim.

Framing in the general versus temporal sense. Some framings may look temporal, but they may not be such. Take *primacy*: in the general sense it means superiority of quality (“*This is the best policy*”), while in the temporal sense, primacy refers to being first in time (“*This was the first policy of its kind*”). Similarly, “*Once a booming hub, now a city in decline*” expresses temporal contrast, whereas “*This economy is worse than others*” conveys a general comparison without temporal framing. Distinguishing these senses is essential to avoid mistaking non-temporal framings for temporal ones.

We provide additional examples for all frames with contrasting cases in Appendix B.

4 Corpus Description

4.1 Article Selection

We construct our corpus through a multi-stage pipeline, beginning with large-scale URL retrieval from GDELT (Leetaru and Schrodt, 2013), followed by successive article extraction, filtering, preprocessing, and sampling steps.

Themes We source article URLs from the GDELT Global Knowledge Graph. We filter records by theme to construct a diverse news corpus spanning political, economic, cultural, commercial, and crisis-related reporting published between September 2024 and September 2025.

Corpus Retrieval We restrict retrieval to a set of widely circulated English and German news outlets, yielding ~ 2 M candidate articles. From this pool, we sample 6,000 article URLs using stratification within each ⟨language, outlet, month⟩ stratum to avoid source imbalance and short-term news bursts.

Text Extraction From the 6,000 URLs, we extract full article text using Trafilatura (Barbarese, 2021), which captures both the article content and associated metadata. We discard failed or garbled extractions and retain only articles containing at least 200 words. We further remove exact duplicates using URL matching and normalized-text hashing, followed by near-duplicate filtering via locality-sensitive hashing (Broder, 1997).

Article Tagging and Sampling An initial annotation pass indicated that opinionated content constituted a minority of the raw news corpus. To increase coverage of persuasive temporal framing, we use LLM-assisted document-level tagging to identify likely opinion articles using GPT-4o (OpenAI, 2024). We combine LLM-guided upsampling with stratified uniform sampling where we target an overall opinionated document proportion of 70% in the final dataset. We retain a substantial portion of non-opinionated articles to account for potential noise in the LLM-generated tags. It is important to highlight that these LLM tags guided sampling only at the document-level; all sentence-level temporal framing labels were assigned by human annotators. We report the final outlet-level sampling distribution in Table 1. We further analyze the prevalence of sentence-level human annotations in articles tagged as opinionated versus non-opinionated by the LLM, as described in Appendix C.1.

Lang	Media Outlet	n	n_{opinion}	n_{sampled}
English	Forbes	397	271	53
	Breitbart	224	80	29
	The Guardian	385	119	28
	Daily Mail	882	108	27
	CNN	356	33	18
	Zero Hedge	104	70	17
	BBC	404	17	11
	Fox News	243	27	10
	China Daily (EN)	148	26	9
	CBS News	100	7	7
	SAMAA TV	57	5	6
	Daily Pakistan (EN)	42	5	6
	NBC News	109	6	5
	Time	28	10	3
	The Epoch Times	175	25	2
	Global Times (EN)	32	4	2
	Liberty Times Net (EN)	17	0	2
	Associated Press	10	2	1
	USA Today	2	1	1
	Kronen Zeitung (EN)	1	0	1
German	Focus Online	107	31	31
	t-online	242	18	28
	Die Welt	359	12	26
	n-tv	276	12	24
	Süddeutsche Zeitung	174	35	22
	Tages-Anzeiger	78	19	20
	Stern	87	12	15
	Bild	125	13	14
	Der Tagesspiegel	109	15	13
	NachDenkSeiten	13	13	11
	Kronen Zeitung	44	4	5
	Epoch Times (DE)	59	1	4
	Frankfurter Allgemeine Zeitung	70	28	3
	Junge Freiheit	5	4	3
FAZ Kaufkompass	4	4	1	

Table 1: **Media outlet-level coverage and final sampling counts.** n = successfully extracted articles. n_{opinion} denotes LLM-tagged indicators of opinionated content, used for sampling only. n_{sampled} denotes articles retained after stratified uniform sampling over time, with upsampling of opinionated articles.

Sentence Segmentation Finally, we segment the articles into sentences with Stanza (Qi et al., 2020) using language-specific models. We prepend titles as the first sentence if they were not already part of the body. After segmentation, we discard fragments consisting only of punctuation as well as fragments containing a single character.

4.2 Annotation Process

We design an iterative annotation workflow with clear guidelines and weekly calibration sessions to ensure consistency across annotators and over time. We use the INCEPTION annotation platform (Klie et al., 2018), which supports span-based and multi-label annotation as well as adjudication and collaborative review. Appendix H shows an example of the annotation interface. Through repeated discussion of disagreements and refinement of guidelines, this process ensured convergence on consistent interpretations of temporal framing, improved annotator alignment, and reduced ambiguity in challenging boundary cases.

Lang.	N	α_{det}	$\bar{\alpha}_{\text{frame}}$	min	max	Set-F1 ₊
EN	1,934	0.605	0.494	0.416	0.590	0.402
DE	2,317	0.720	0.588	0.302	0.768	0.443

Table 2: **Sentence-level inter-annotator agreement.** α_{det} denotes Krippendorff’s α (nominal) for binary detection of any temporal framing. $\bar{\alpha}_{\text{frame}}$ is the mean per-frame α across the eight temporal frames. Set-F1₊ is computed on the positive-only subset and averaged over the sentences.

A coordinator oversaw curation and consistency for each language. The team comprised three English and four German annotators. Annotators met weekly to review disagreements and revised the taxonomy and guidelines. The initial taxonomy included two frames, *momentum* and *cyclicity*, later merged into *continuity* after annotators noted it subsumes both. We updated guidelines to exclude quoted and indirect speech, as these cases caused disagreement over whether the persuasive voice belonged to the author or quoted source. We restricted cross-sentence framing to the anchor sentence, particularly for *temporal contrast*, since annotators observed that persuasive force arises from the anchor sentence. Finally, we broadened *temporal anchoring* from references to historical landmarks to rhetorical anchoring to past events, as defining such landmarks proved subjective and introduced bias, whereas rhetorical past references offered a more consistent criterion.

Inter-Annotator Agreement. We randomly assign ~ 50 articles per language to 2–4 annotators. At the sentence level (Table 2), Krippendorff’s α (nominal) for binary detection of temporal framing (presence vs. absence) reaches 0.605 for English and 0.720 for German. These agreement levels are consistent with prior work on fine-grained framing and rhetorical annotation, where moderate agreement is typical due to the interpretive and multi-label nature of the task (Card et al., 2015; Piskorski et al., 2023). We also compute per-frame agreement by treating each of the eight temporal frames as a binary label, yielding mean α values of 0.494 (EN) and 0.588 (DE). To account for the multi-label nature of frame assignment, we further report positive-only set agreement, with Set-F1₊ scores of 0.402 (EN) and 0.443 (DE).

Appendix C.2 reports per-frame sentence-level agreement scores, and Appendix C.3 presents *document-level* agreement analysis.

Lang.	Docs	Sents	Words	Chars	AVG _s	AVG _w	AVG _c	SentAnn	Ann	AVG _s	AVG _a
EN	238	8,750	161,311	992,556	36.76	677.78	4,170.40	1,748	2,334	7.34	1.34
DE	220	11,909	163,381	1,147,357	54.13	742.64	5,215.26	617	696	2.80	1.13
Total	458	20,659	324,692	2,139,913	45.11	708.93	4,672.30	2,365	3,030	5.16	1.28

Table 3: **Corpus statistics.** The table shows the total number of documents (Docs), sentences (Sents), words (Words), and characters (Chars) by language. Averages per document are provided for sentences (AVG_s), words (AVG_w), and characters (AVG_c). The table also reports the number of sentences containing at least one annotation (SentAnn), the total number of annotations (Ann), the average number of annotated sentences per document (AVG_s), and the average number of annotations per annotated sentence (AVG_a).

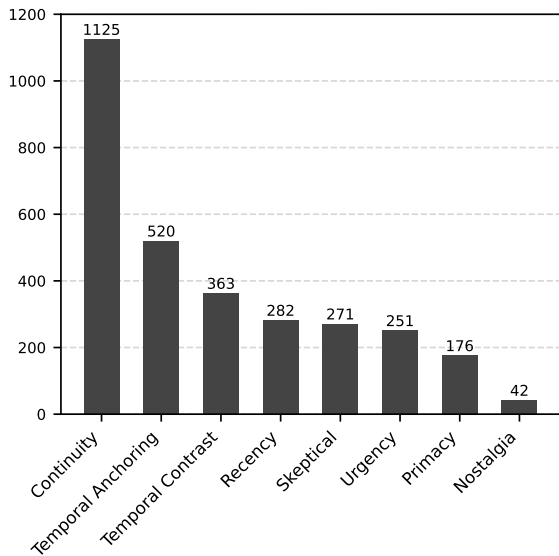


Figure 2: **Distribution of frames in the corpus.**

4.3 Corpus Analysis

4.3.1 Statistics

Table 3 presents the overall corpus statistics with a breakdown by language. While English and German documents are comparable in size, annotations are more concentrated in English, with a higher proportion of sentences labeled as temporally framed. Moreover, temporally framed English sentences contain slightly more temporal frames on average than German. Figure 2 shows the distribution of temporal frames across the dataset. *Continuity*, *Temporal Anchoring*, and *Temporal Contrast* are the most frequent frames, indicating a dominant rhetorical focus on linking past and present or emphasizing temporal change. The prominence of *Continuity* is particularly notable, as it reflects the tendency of news discourse to frame events as part of ongoing trajectories, whether to legitimize claims via persistence, normalize repetition, or express dissatisfaction with an unchanging situation.

Co-occurrence of Frames Most temporally framed sentences contain a single frame (1,803 sentences), with progressively fewer sentences exhibiting multiple frames (468 with two frames, 86 with three, 7 with four, and 1 with five). To understand these cases, we examine which temporal frames appear together (see Appendix C.4). *Temporal Anchoring* co-occurs with *Continuity*, *Primacy*, and *Temporal Contrast*. This is natural, as anchoring situates claims in time by reinforcing persistence, setting precedence via an anchor point, or highlighting change. *Skeptical* framing shows moderate overlap with *Urgency*, often in contexts where caution and doubt are raised about the future.

Corpus Splits To control for split-specific bias, we construct the train, development, and test splits such that each split contains a comparable proportion of documents with and without temporal framing *within each language*. While the overall framing prevalence differs across languages, the splits preserve each language’s internal framing distribution across documents. The dataset follows an approximate 70/13/17 train-development-test distribution. See Appendix C.5 for a detailed breakdown across splits.

Lexical Analysis To characterize how temporal frames are realized at the surface level, we conduct a lexical association analysis comparing sentences annotated with each temporal frame against the rest of the corpus using smoothed log-odds ratios (Monroe et al., 2008). Overall, most frames exhibit intuitive lexical cues, while others rely less on stable surface forms. In particular, *Temporal Anchoring* and *Nostalgia* show weaker and more heterogeneous lexical signals resulting from their dependence on event-specific references and evaluative language rather than recurring temporal expressions. A detailed lexical analysis, including per-frame log-odds profiles and diagnostics, is provided in Appendix C.6.

Method	Model	Binary Detection			Multilabel Classification			
		P	R	F1	P	R	Micro-F1	Macro-F1
	Random Baseline	0.11	0.88	0.20	0.02	0.49	0.04	0.03
Zero-Shot	Llama-3.1-8B-Instruct	0.14	0.94	0.24	0.06	0.34	0.10	0.11
	Llama-3.3-70B-Instruct	0.28	0.78	0.41	0.16	0.43	0.23	0.21
	Qwen3-8B	0.21	0.75	0.33	0.08	0.29	0.13	0.14
	Qwen3-32B	0.36	0.57	0.44	0.20	0.31	0.25	0.22
	Qwen3-235B-A22B	0.31	0.70	0.44	0.17	0.38	0.24	0.23
	GPT-5.2	0.34	0.67	0.45	0.24	0.46	0.31	0.29
Supervised	XLm-R	0.38	0.75	0.51	0.28	0.57	0.37	0.33
	Llama-3.1-8B-Instruct	0.67	0.45	0.54	0.48	0.37	0.42	0.35
	Qwen3-8B	0.60	0.54	0.57	0.48	0.40	0.44	0.35

Table 4: **Overall results on the test set.** Zero-shot models show high recall but low precision, while supervised models achieve more balanced performance with strong gains in detection and multilabel classification. The random baseline samples labels from the annotation label distributions in the combine training and development sets.

5 Experiments

5.1 Experimental Setup

We evaluate two strategies for sentence-level temporal frame detection: (1) zero-shot classification and (2) supervised fine-tuning on our annotated corpus. The former tests out-of-the-box recognition without task-specific training, while the latter assesses learning under supervision, using a consistent input representation.

We adopt a sentence-only input for all experiments, motivated by context ablation results in Appendix E.2.1. We also evaluate two extended variants: (i) neighboring context (preceding and following sentences) and (ii) full document context. Neither improves performance over sentence-only input and incurs higher token costs. Additional context makes models more conservative, often consistently leading to missed temporal frames in the target sentence during evaluation.

Zero-Shot Classification. We use both a proprietary LLM and open-weight language models of various scales for zero-shot inference. Specifically, we evaluate OpenAI’s GPT-5.2 (OpenAI, 2025) alongside models from the LLaMA3 (Llama Team, 2024) and Qwen3 (Qwen Team, 2025) architecture families. For the open-weight models, we consider sizes ranging from 8 to 235 billion parameters to examine the effect of scale on performance. We prompt all zero-shot models using a consistent instruction set (see Appendix G), with additional inference and output-schema details provided in Appendix E.2.2.

Supervised Fine-Tuning. For supervised fine-tuning, we consider both an efficient encoder-based model and selected language models from our zero-shot experiments. We adopt the pre-trained encoder model XLm-RoBERTa-base (XLm-R) (Conneau et al., 2020), which has 270 million parameters, as an efficient baseline and train it as a multi-label sentence classification model. Given the extreme class imbalance in the data, where most sentences contain no temporal frame and only a small fraction exhibit multiple frames, we optimize a weighted binary cross-entropy loss that down-weights negative instances and apply label-aware batching to reduce batch-level sparsity and stabilize training (see Appendix E.1 for details). Beyond the encoder-based baseline, we fine-tune two representative language models from our zero-shot selection, namely Qwen3-8B and LLaMA3.1-8B. Details of the fine-tuning setup for these models are provided in Appendix E.2.3.

5.2 Results

Table 4 summarizes results on the test set, which contains 3,517 sentences. We evaluate zero-shot configurations over ten runs and fine-tuned models over a single run, reporting mean scores across all metrics. We report both binary detection and multilabel classification performance. Figure 3 visualizes Micro-F1 as a function of model size.

Among *zero-shot approaches*, GPT-5.2 achieves the strongest performance, with a binary detection F1 of 0.45 and a multilabel Micro-F1 of 0.31. Within the Qwen family, binary F1 increases from 0.33 for Qwen3-8B to 0.44 for Qwen3-235B.

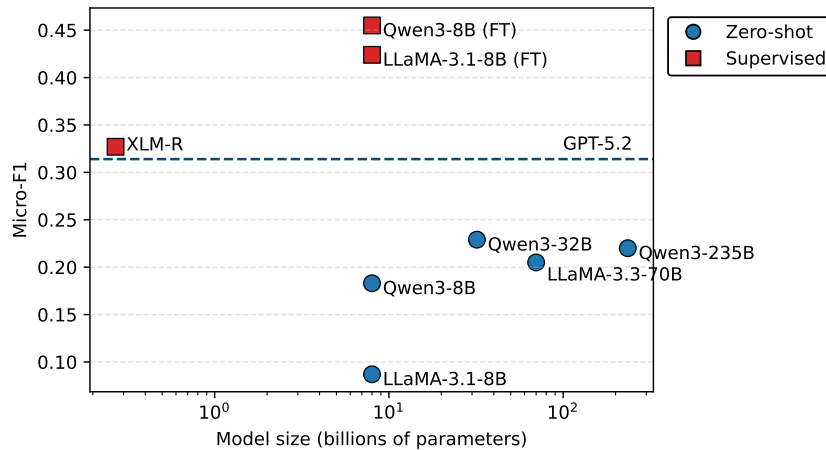


Figure 3: **Micro-F1 vs. model size on the test set.** Supervised fine-tuning (red squares) substantially outperforms zero-shot inference (blue circles) across all model scales. XLM-RoBERTa achieves competitive performance with only 0.27B parameters.

A similar trend is observed for LLaMA models, where the 70B variant (0.41) substantially outperforms the 8B model (0.24), highlighting the effect of increased model capacity and representational power. Across zero-shot models, recall is generally high while precision remains low, indicating a tendency toward over-prediction and reduced reliability in practical settings and downstream applications. This suggests that, without task-specific supervision, LLMs struggle to reliably detect temporally framed sentences.

Supervision yields substantial improvements over zero-shot baselines. XLM-R achieves a binary F1 of 0.51 and a Micro-F1 of 0.37 with only 270M parameters, which demonstrates that encoder-based models can effectively learn temporal framing patterns under direct supervision. For LLMs in the zero-shot setting, scale alone produces only modest gains, with improvements tapering as model size increases, suggesting diminishing returns from additional parameters and limited benefits from scaling alone (Figure 3). In contrast, fine-tuning leads to large performance gains: LLaMA-3.1-8B-Instruct improves by 125% relative to its zero-shot counterpart, while fine-tuned Qwen3-8B achieves *the best overall performance* with a binary F1 of 0.57 and a Micro-F1 of 0.44, corresponding to a 27% relative improvement over the strongest zero-shot result. The fine-tuned LLMs also exhibit distinct precision–recall profiles: Qwen3-8B remains relatively balanced, whereas LLaMA-3.1-8B-Instruct is more conservative, achieving higher precision at the cost of lower recall across different evaluation scenarios and dataset conditions.

5.3 Discussion

Learnability and scale limitations. Our results (Figure 3) indicate that sentence-level temporal framing is learnable, but not reliably addressed by model scale alone in the zero-shot setting. While larger models yield incremental gains, performance saturates quickly, suggesting that temporal framing relies on subtle rhetorical cues and relational contrasts that benefit more from task-specific supervision than from additional parameters. This indicates that prediction errors are driven by label ambiguity, subtle rhetorical cues, and sparsity rather than raw model capacity, highlighting the importance of high-quality annotations and task-specific training signals. Supervised training consistently improves both binary detection and frame identification, supporting the view that temporal framing patterns are systematic but not easily recoverable through general-purpose prompting alone.

Frame-level behavior. At a finer granularity, supplementary analysis in Appendix E.3 provides per-language and per-frame breakdowns that further contextualize these findings and reveal cross-lingual variation. Across zero-shot models, frames such as *Continuity*, *Temporal Contrast*, and *Urgency* are more consistently detected, likely because they are more prevalent and associated with recurring lexical or structural cues (see Appendix C.6). Supervised fine-tuning reduces this gap across most frames, improving robustness and consistency, though rare frames, particularly *Nostalgia*, remain affected by data sparsity and limited training instances.

Cross-lingual and sparsity analysis. Across languages, performance differs: per-language analysis shows lower performance on German data relative to English, which is expected given that the German subset contains substantially fewer framed sentences and fewer positive instances per frame, as discussed in Section 4.3.1. This sparsity disproportionately affects recall for low-frequency frames and contributes to greater variance across models, rather than indicating a failure of cross-lingual generalization. Future work that explicitly controls for sparsity and corpus composition, for example through targeted augmentation of rare frames, would help disentangle language from data distribution effects.

Model behavior and trade-offs. Beyond these data effects, encoder-based and instruction-tuned models exhibit complementary strengths. XLM-R shows that a standard multilingual encoder, when trained with imbalance-aware objectives, can capture much of the underlying signal, achieving strong performance despite its relatively small size and surpassing all zero-shot baselines. In contrast, fine-tuned 8B LLMs deliver the strongest end-to-end performance under the same sentence-only input format. At the same time, they exhibit distinct precision–recall profiles: LLaMA-3.1-8B-Instruct is more conservative, favoring higher precision, whereas Qwen3-8B maintains a more balanced recall profile. These differences make the models better suited to different downstream settings, depending on whether minimizing false positives or maximizing coverage is the primary objective.

6 Conclusion and Future Work

We introduced *temporal framing* as a distinct and computationally tractable dimension of persuasive language. We formalized temporal framing as the rhetorical use of time-related expressions to structure meaning rather than to report chronology, proposed a taxonomy of eight temporal frames grounded in social science theory, and realized it through a multilingual, sentence-level annotated news corpus in English and German. Our corpus analysis showed that temporal framing is frequent, heterogeneous, and often multi-faceted, with frames such as *Continuity*, *Temporal Anchoring*, and *Temporal Contrast* playing a central role in news discourse and shaping how narratives evolve across different topics and contexts over time and across diverse media outlets.

Our experiments show that temporal framing is *learnable at the sentence level*, but not reliably captured by model scale alone in zero-shot settings. Larger models yield only modest gains, with performance quickly saturating without supervision. In contrast, supervised fine-tuning substantially improves both detection and frame assignment, even for compact models, indicating that temporal framing patterns are systematic but require task-specific learning signals. Fine-tuned models also exhibit distinct precision–recall profiles: Qwen3-8B is more balanced, while LLaMA-3.1-8B-Instruct is more conservative, achieving higher precision at the cost of lower recall.

These findings showed that temporal framing introduced a level of abstraction beyond conventional temporal analysis in NLP. Unlike tasks focused on temporal expression extraction or event ordering, temporal framing required interpreting how time is used to guide interpretation and influence perception. This helped explain why zero-shot models struggled despite strong general language understanding, as the task involved recognizing rhetorical intent rather than surface patterns alone. At the same time, consistent gains from supervision indicated that these patterns are learnable when appropriate signals are provided.

This work provided a valuable resource and empirical foundation for studying the rhetorical functions of time in news discourse. While our work focused on sentence-level modeling, the corpus was designed to support broader analyses at multiple levels of discourse, including document-level analysis. Future work can build on this corpus to examine temporal framing across additional languages and domains, to model frame interactions beyond the sentence level, and to investigate how temporal framing shapes narrative structure at scale.

Finally, our results highlighted that temporal framing occupies a distinct space between surface-level temporal expression and higher-level narrative structure. While models can learn to detect framing at the sentence level, many cases require interpreting implicit rhetorical cues that extend beyond explicit temporal markers. Thus, future work should explore richer representations of discourse context, including cross-sentence dependencies and narrative progression, to better capture how temporal framing unfolds across longer text spans. Such extensions would enable a more complete account of how time is used not only to describe events, but to shape interpretation and persuasion.

Limitations

Corpus The corpus comprises 458 English and German news articles from a fixed set of outlets. It is enriched for opinionated content to increase temporal framing coverage and is not representative of overall news production. The corpus is unbalanced across outlets and framing prevalence, with variation across languages. As it is limited to English and German news, findings may not generalize to other languages or media contexts.

Annotation Scope Annotations are performed at the sentence level rather than using span-level annotation, which may miss cases where temporal framing is realized within partial clauses or spans that do not align with sentence boundaries. In addition, annotations focus on authorial voice and exclude quoted and indirect speech, limiting coverage of framing conveyed through attribution.

Ethical Considerations

Biases The corpus draws on a fixed set of news outlets and reflects their inherent biases. Although it includes diverse editorial orientations, it is not balanced across viewpoints or ideologies. Annotators were instructed to identify temporal framing independently of personal views or article stance.

Intended Use and Misuse Potential The corpus supports research on temporal framing and rhetorical analysis in news, aiding transparency and critical understanding of persuasive language. However, as with other framing resources, annotations could be misused to optimize persuasive or manipulative communication. The corpus is released for analytical purposes with awareness of this risk.

Annotation Practice and Fairness. Annotations were produced by trained annotators affiliated with the authors' institutions, following detailed guidelines with periodic calibration. Annotators were compensated according to institutional standards, and no crowd-sourcing was used. The task does not assess factual accuracy, truthfulness, or political intent, and labels should not be interpreted as indicators of misinformation or bias.

Environmental Considerations This work makes use of large language models for inference. As with other LLM-based methods, this entails computational cost and contributes to energy consumption and associated carbon emissions.

Acknowledgments

This work was partially supported by the German Federal Ministry for Research, Technology and Space (BMFTR) under projects “VeraXtract” (16IS24066) and “news-polygraph” (03RU2U151C).

References

- George Ainslie. 1975. [Specious reward: a behavioral theory of impulsiveness and impulse control](#). *Psychological Bulletin*, 82(4):463.
- Stuart Albert. 1977. [Temporal comparison theory](#). *Psychological Review*, 84(6):485.
- David Alonso del Barrio and Daniel Gatica-Perez. 2023. [Framing the news: From human perception to large language model inferences](#). In *Proceedings of the 2023 ACM International Conference on Multimedia Retrieval, ICMR '23*, page 627–635, New York, NY, USA. Association for Computing Machinery.
- Solomon E Asch. 1946. [Forming impressions of personality](#). *The Journal of Abnormal and Social Psychology*, 41(3):258.
- Adrien Barbaresi. 2021. [Trafilatura: A web scraping library and command-line tool for text discovery and extraction](#). In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing: System Demonstrations*, pages 122–131, Online. Association for Computational Linguistics.
- A.Z. Broder. 1997. [On the resemblance and containment of documents](#). In *Proceedings. Compression and Complexity of SEQUENCES 1997 (Cat. No.97TB100171)*, pages 21–29.
- Claus-Christian Carbon and Sabine Albrecht. 2012. [Bartlett's schema theory: the unreplicated “portrait d’homme” series from 1932](#). *Quarterly Journal of Experimental Psychology*, 65(11):2258–2270.
- Dallas Card, Amber E. Boydston, Justin H. Gross, Philip Resnik, and Noah A. Smith. 2015. [The media frames corpus: Annotations of frames across issues](#). In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 438–444, Beijing, China. Association for Computational Linguistics.
- Dallas Card, Serina Chang, Chris Becker, Julia Mendelsohn, Rob Voigt, Leah Boustan, Ran Abramitzky, and Dan Jurafsky. 2022. [Computational analysis of 140 years of US political speeches reveals more positive but increasingly polarized framing of immigration](#). *Proceedings of the National Academy of Sciences*, 119(31):e2120510119.

- Sucharita Chandran and Geeta Menon. 2004. [When a day means more than a year: Effects of temporal framing on judgments of health risk](#). *Journal of Consumer Research*, 31(2):375–389.
- Zheng Chu, Jingchang Chen, Qianglong Chen, Weijiang Yu, Haotian Wang, Ming Liu, and Bing Qin. 2024. [TimeBench: A comprehensive evaluation of temporal reasoning abilities in large language models](#). In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1204–1228, Bangkok, Thailand. Association for Computational Linguistics.
- Robert B Cialdini. 2001. [The science of persuasion](#). *Scientific American*, 284(2):76–81.
- Manuel Jesús Cobo, Alberto Maria De Mascellis, Michelangelo Misuraca, and Maria Spano. 2025. [Automatic temporal framing detection for news analysis](#). *Quality & Quantity*.
- Bernard Comrie. 1985. *Tense*. Cambridge Textbooks in Linguistics. Cambridge University Press, Cambridge.
- Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2020. [Unsupervised cross-lingual representation learning at scale](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8440–8451, Online. Association for Computational Linguistics.
- Rohan Das, Aditya Chandra, I-Ta Lee, and Maria Leonor Pacheco. 2024. [Media framing through the lens of event-centric narratives](#). In *Proceedings of the 6th Workshop on Narrative Understanding*, pages 85–98, Miami, Florida, USA. Association for Computational Linguistics.
- Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, and Luke Zettlemoyer. 2023. [QLoRA: Efficient finetuning of quantized LLMs](#). In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS '23, Red Hook, NY, USA. Curran Associates Inc.
- Xi Ding and Lei Wang. 2025. [Do language models understand time?](#) In *Companion Proceedings of the ACM on Web Conference 2025*, WWW '25, page 1855–1868, New York, NY, USA. Association for Computing Machinery.
- Nicholas Evans and Stephen C. Levinson. 2009. [The myth of language universals: Language diversity and its importance for cognitive science](#). *Behavioral and Brain Sciences*, 32(5):429–492.
- Edward A. Feigenbaum and Herbert A. Simon. 1962. [A theory of the serial position effect](#). *British Journal of Psychology*, 53(3):307–320.
- Craig R Fox and Amos Tversky. 1995. [Ambiguity aversion and comparative ignorance](#). *The Quarterly Journal of Economics*, 110(3):585–603.
- Adrian Furnham and Hua Chu Boo. 2011. [A literature review of the anchoring effect](#). *The Journal of Socio-Economics*, 40(1):35–42.
- Kai Gehring and Matteo Grigoletto. 2023. [Analyzing climate change policy narratives with the character-role narrative framework](#). CESifo Working Paper 10429, CESifo.
- Maurice Halbwachs. 1950. *La mémoire collective*. Presses Universitaires de France.
- François Hartog and Saskia Brown. 2015. *Regimes of Historicity: Presentism and Experiences of Time*. Columbia University Press.
- Duygu Sezen Islakoglu and Jan-Christoph Kalo. 2025. [ChronoSense: Exploring temporal understanding in large language models with time intervals of events](#). In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 590–602, Vienna, Austria. Association for Computational Linguistics.
- John T Jost and Mahzarin R Banaji. 1994. [The role of stereotyping in system-justification and the production of false consciousness](#). *British Journal of Social Psychology*, 33(1):1–27.
- Daniel Kahneman and Amos Tversky. 1979. [Prospect theory: An analysis of decision under risk](#). *Econometrica*, 47(2):263–291.
- David E. Kanouse and L. Reid Hanson Jr. 1987. Negativity in evaluations. In Edward E. Jones, David E. Kanouse, Harold H. Kelley, Richard E. Nisbett, Stuart Valins, and Bernard Weiner, editors, *Attribution: Perceiving the Causes of Behavior*, pages 47–62. Lawrence Erlbaum Associates.
- Kyongseok Kim and Hyang-Sook Kim. 2018. [Time matters: Framing antismoking messages using current smokers' preexisting perceptions of temporal distance to smoking-related health risks](#). *Health Communication*, 33(3):338–348.
- Jan-Christoph Klie, Michael Bugert, Beto Boullosa, Richard Eckart de Castilho, and Iryna Gurevych. 2018. [The INCEpTION platform: Machine-assisted and knowledge-oriented interactive annotation](#). In *Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations*, pages 5–9, Santa Fe, New Mexico. Association for Computational Linguistics.
- Woosuk Kwon, Zhuohan Li, Siyuan Zhuang, Ying Sheng, Lianmin Zheng, Cody Hao Yu, Joseph Gonzalez, Hao Zhang, and Ion Stoica. 2023. [Efficient memory management for large language model serving with PagedAttention](#). In *Proceedings of the 29th Symposium on Operating Systems Principles*, SOSP '23, page 611–626, New York, NY, USA. Association for Computing Machinery.

- Davide Lamorte, Marco Rovera, Alfio Ferrara, and Sara Tonelli. 2024. [TimeFrame: Querying and visualizing event semantic frames in time](#). In *Proceedings of the First Workshop on Reference, Framing, and Perspective @ LREC-COLING 2024*, pages 13–17, Torino, Italia. ELRA and ICCL.
- Kalev Leetaru and Philip A. Schrod. 2013. [GDELT: Global data on events, location, and tone](#).
- Siyi Liu, Lei Guo, Kate Mays, Margrit Betke, and Derry Tanti Wijaya. 2019. [Detecting frames in news headlines and its application to analyzing news framing trends surrounding U.S. gun violence](#). In *Proceedings of the 23rd Conference on Computational Natural Language Learning (CoNLL)*, pages 504–514, Hong Kong, China. Association for Computational Linguistics.
- Llama Team. 2024. [The Llama 3 Herd of Models](#). *CoRR*, abs/2407.21783.
- Tarek Mahmoud, Zhuohan Xie, Dimitar Iliyanov Dimitrov, Nikolaos Nikolaidis, Purificação Silvano, Roman Yangarber, Shivam Sharma, Elisa Sartori, Nicolas Stefanovitch, Giovanni Da San Martino, Jakub Piskorski, and Preslav Nakov. 2025. [Entity framing and role portrayal in the news](#). In *Findings of the Association for Computational Linguistics: ACL 2025*, pages 302–326, Vienna, Austria. Association for Computational Linguistics.
- William J McGuire. 1960. [Cognitive consistency and attitude change](#). *The Journal of Abnormal and Social Psychology*, 60(3):345.
- Terence R. Mitchell, Leigh Thompson, Erika Peterson, and Randy Cronk. 1997. [Temporal adjustments in the evaluation of events: The “rosy view”](#). *Journal of Experimental Social Psychology*, 33(4):421–448.
- Burt L. Monroe, Michael P. Colaresi, and Kevin M. Quinn. 2008. [Fightin’ words: Lexical feature selection and evaluation for identifying the content of political conflict](#). *Political Analysis*, 16(4):372–403.
- Bennet B Murdock Jr. 1962. [The serial position effect of free recall](#). *Journal of Experimental Psychology*, 64(5):482.
- Ted O’Donoghue and Matthew Rabin. 2015. [Present bias: Lessons learned and to be learned](#). *American Economic Review*, 105(5):273–279.
- OpenAI. 2024. [GPT-4o system card](#). Technical report.
- OpenAI. 2025. [GPT-5.2 system card](#). Technical report.
- Sheina Orbell and Maria Kyriakaki. 2008. [Temporal framing and persuasion to adopt preventive health behavior: Moderating effects of individual differences in consideration of future consequences on sunscreen use](#). *Health Psychology*, 27(6):770–779.
- Yulia Otmakhova, Shima Khanehzar, and Lea Frermann. 2024. [Media framing: A typology and survey of computational approaches across disciplines](#). In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 15407–15428, Bangkok, Thailand. Association for Computational Linguistics.
- Jakub Piskorski, Nicolas Stefanovitch, Giovanni Da San Martino, and Preslav Nakov. 2023. [SemEval-2023 task 3: Detecting the category, the framing, and the persuasion techniques in online news in a multilingual setup](#). In *Proceedings of the 17th International Workshop on Semantic Evaluation (SemEval-2023)*, pages 2343–2361, Toronto, Canada. Association for Computational Linguistics.
- Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton, and Christopher D. Manning. 2020. [Stanza: A Python natural language processing toolkit for many human languages](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 101–108, Online. Association for Computational Linguistics.
- Qwen Team. 2025. [Qwen3 technical report](#). *CoRR*, abs/2505.09388.
- William Samuelson and Richard Zeckhauser. 1988. [Status quo bias in decision making](#). *Journal of risk and uncertainty*, 1(1):7–59.
- Constantine Sedikides, Tim Wildschut, Jamie Arndt, and Clay Routledge. 2008. [Nostalgia: Past, present, and future](#). *Current Directions in Psychological Science*, 17(5):304–307.
- Constantine Sedikides, Tim Wildschut, and Denise Baden. 2004. [Nostalgia: Conceptual issues and existential functions](#). In Jeff Greenberg, Sander L. Koole, and Tom Pyszczynski, editors, *Handbook of Experimental Existential Psychology*, pages 200–214. The Guilford Press, New York, NY, US.
- John E. Smith. 1969. [Time, times, and the “right time”: Chronos and kairos](#). *The Monist*, 53(1):1–13.
- Veronika Solopova, Oana-Iuliana Popescu, Christoph Benz Müller, and Tim Landgraf. 2023. [Automated multilingual detection of pro-Kremlin propaganda in newspapers and Telegram posts](#). *Datenbank-Spektrum*, 23(1):5–14.
- Dominik Stambach, Maria Antoniak, and Elliott Ash. 2022. [Heroes, villains, and victims, and GPT-3: Automated extraction of character roles without training data](#). In *Proceedings of the 4th Workshop of Narrative Understanding (WNU2022)*, pages 47–56, Seattle, United States. Association for Computational Linguistics.
- Jean Starobinski. 1966. [The idea of nostalgia](#). *Diogenes*, 14(54):81–103.
- Jerry Suls and Ladd Wheeler, editors. 2000. *Handbook of Social Comparison: Theory and Research*. Kluwer Academic Publishers.

- Paul A. Taggart. 2000. *Populism*. Concepts in the social sciences. Open University Press, Buckingham.
- Qingyu Tan, Hwee Tou Ng, and Lidong Bing. 2023. [Towards benchmarking and improving the temporal reasoning capability of large language models](#). In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 14820–14835, Toronto, Canada. Association for Computational Linguistics.
- Yaacov Trope and Nira Liberman. 2010. [Construal-level theory of psychological distance](#). *Psychological Review*, 117(2):440.
- Amos Tversky and Daniel Kahneman. 1973. [Availability: A heuristic for judging frequency and probability](#). *Cognitive Psychology*, 5(2):207–232.
- Amos Tversky and Daniel Kahneman. 1974. [Judgment under uncertainty: Heuristics and biases](#). *Science*, 185(4157):1124–1131.
- Brady Wagoner and Lisa Herbig. 2025. [Thinking through historical analogies: A longitudinal analysis of sense-making during the pandemic](#). *Integrative Psychological and Behavioral Science*, 59(4):68.
- Tim Wildschut, Constantine Sedikides, Jamie Arndt, and Clay Routledge. 2006. [Nostalgia: Content, triggers, functions](#). *Journal of Personality and Social Psychology*, 91(5):975–993.
- Thomas A Wills. 1981. [Downward comparison principles in social psychology](#). *Psychological Bulletin*, 90(2):245.
- Anne E Wilson and Michael Ross. 2001. [From chump to champ: people’s appraisals of their earlier and present selves](#). *Journal of Personality and Social Psychology*, 80(4):572.
- Ruth Wodak. 2015. *The Politics of Fear: What Right-Wing Populist Discourses Mean*. SAGE.

A Annotation Guidelines

The annotation for this task is done as follows:

1. You should carefully read Section 3 before beginning annotation, in order to understand the definition of temporal framing, the taxonomy of frames, and key distinctions such as factual vs. opinionated use and general vs. temporal senses.
 2. You are provided with a number of news articles and are expected to identify sentences that contain temporal framing according to the taxonomy introduced in Figure 1b.
 3. The taxonomy is provided in two forms:
 - **Brief overview** (Figure 1b): concise definitions and examples for quick reference.
 - **Detailed taxonomy description** (Section B): extended definitions, rhetorical functions, positive as well as negative examples, and explanations for each example. The detailed description should be consulted for challenging and ambiguous cases.
 4. You should **only mark positive instances of framing**. Negative examples in the taxonomy are included to illustrate potential pitfalls and help avoid incorrect annotations.
 5. A sentence should be annotated as a frame only if the **temporal element is central to its persuasive force**. If the temporal element is merely descriptive (factual reporting) or absent (non-temporal lookalike), then no frame should be assigned.
 6. You should rely only on the information in the text itself, not on external world knowledge or assumptions.
 7. When in doubt, apply the following heuristic: *if removing the temporal expression leaves the persuasive force of the sentence unchanged, then it is not an instance of temporal framing*. For example, consider: “The plan was announced last week and it is deeply flawed.”
 - With the temporal expression: the persuasive force comes from the claim that the plan is *deeply flawed*.
 - Without the temporal expression (“The plan is deeply flawed”): the persuasive force remains identical.
- Thus, the temporal cue “last week” is incidental, and this sentence should not be considered for any temporal framing.
8. Pay attention to common temporal cues, including but not limited to: *first, latest, now, then, still, again, since, until, before, after, within, deadline, turning point, every, seasonal, for years/decades/centuries, post-9/11, pre-war, next year, last month*. Such cues may signal the existence of a temporal frame.
 9. Some sentences may include more than one temporal frame simultaneously (e.g., *recency + urgency*). In such cases, you should assign all frames that apply.
 10. Context matters. A sentence that appears factual in isolation may still count as framing if the surrounding context utilizes its temporal component as a persuasive element.
 11. **Quoted or Indirect Speech Exclusion:** Do not annotate any sentence that is wholly within quotation marks or represents indirect or reported speech (e.g., “The minister said the crisis will worsen next year”). Temporal framing analysis applies only to the narrator’s or authorial voice, not to speech attributed to others.
- Example:**
- “Back then, I simply was not as educated as to the implications and consequences of my actions as I have become since,” Knight said.
- Annotation: None: indirect or quoted speech is excluded from analysis.*
- Note: Had this not been a quotation, the contrast between “back then” and “since” would qualify as temporal contrast, as it juxtaposes a past state of ignorance with a present state of awareness.*
12. **Cross-Sentence Temporal Framing** Temporal framing may rely on information introduced in earlier sentences. In such cases, **annotate only the sentence where the temporal framing is made explicit or persuasive,**

even if earlier sentences provide necessary context.

Example:

S_1 : For decades, the agency was trapped in slow, paper-based workflows.

S_2 : Today, it operates digitally, transforming weeks of delays into same-day approvals.

Annotation: S_2 : **Temporal Contrast** (past vs. present).

Recommended workflow.

1. Read the sentence carefully, keeping the surrounding context in mind, and note any temporal cues, both explicit (e.g., “today,” “last year,” “within weeks”) and implicit (e.g., historical landmarks such as “the Cold War era,” “Brooklyn’s protests”).
2. **First check: Quoted/indirect speech.** If the sentence is wholly or primarily direct or indirect speech, do not annotate and skip. Consider for annotation only if the author’s/narrator’s sentence independently carries the temporal framing.
3. **Second check: Fact vs. rhetorical.** Decide whether the temporal element is merely *descriptive* (factual reporting) or functions as *rhetorical framing* (adds persuasive force). If the sentence appears factual in isolation but surrounding context makes the temporal element rhetorically persuasive, then treat it as temporal framing.
4. **Third check: General vs. temporal sense.** If the language only resembles framing in a general sense (e.g., contrast) but does not involve time as a framing resource, treat it as a non-temporal lookalike and do not annotate.
5. **Fourth check: Cross-sentence (anchor-only).** If the sentence contains an incomplete temporal cue (*now, still, no longer, has since, used to*) or implies a then-now contrast, scan the context for potential anchors. Annotate *only* the sentence where the persuasive temporal force culminates as the *anchor*. If the force culminates in a neighboring sentence, skip the current one and annotate that neighbor as the anchor.

6. If the instance passes all checks, select the most appropriate frame(s) from the taxonomy.
7. In cases where multiple frames apply simultaneously, assign all that are relevant.
8. Use the compact taxonomy table for quick reference; consult the detailed taxonomy descriptions if uncertain.
9. **Challenging examples.** If an example remains ambiguous even after consulting the taxonomy, you should note it down and raise it to the language coordinator for resolution.

B Detailed Taxonomy Description

In what follows, we present each temporal frame with a more detailed description, including its *persuasive function*, and examples that separate persuasive uses from mere factual reports and from non-temporal framing.

Note: Examples illustrate how statements may function as temporal framing devices. In some cases, the same wording could be read as neutral description when taken in isolation, but operate as framing when used rhetorically in context. The labels here indicate the intended illustrative function rather than an absolute classification.

B.1 Primacy

Definition: *Primacy framing assigns significance to temporal precedence, treating leading positions as inherently meaningful.*

The precedence may be absolute in time (e.g., “the earliest case ever” or “first since the beginning of the record”) or relative to a bounded interval (e.g., “the earliest this year,” “first in this decade,” or “first under this administration”).

Rhetorical function: *It converts the mere chronology of being first in time into persuasive force, establishing authority, legitimacy, or caution.*

✓ **“As the first city to adopt this policy, we set the standard for the region.”**

Being first is explicitly linked to leadership and legitimacy, making temporal precedence the persuasive force.

✓ **“The collapse of the bank unsettled confidence in the markets to unprecedented levels.”**

Low market confidence is framed as without

precedent, amplifying the perceived danger of the bank's collapse.

X “The vaccine was approved first.”

Taken in isolation, this is a chronological report; it does not claim that being first entails special significance. However, if being first plays a persuasive role when the surrounding context is taken into account, then the same sentence could function as primacy framing.

X “The best-performing vaccine won praise.”

The praise is grounded in quality (performance), not temporal primacy; it resembles primacy in a general sense but lacks the temporal element.

B.2 Recency

Definition: *Recency framing attributes significance to temporal proximity, treating the latest events as inherently meaningful.*

The proximity may be absolute in time (e.g., “the most recent ever,” “latest on record”), relative to a bounded interval (e.g., “latest this year,” “most recent this quarter,” “latest under this administration”), or anchored to a salient event (e.g., “just hours after the event” “a few days after the cease-fire”) where closeness to that event is presented as meaningful.

Rhetorical function: *It seizes attention through temporal proximity, subtly displaces older evidence without needing to disprove it, or pressures audiences to accept arguments as timely and relevant.*

✓ “Today’s figures matter most for judging performance.”

Freshness is explicitly framed as the reason for relevance.

✓ “Just released footage shows what really happened.”

The recency of the footage is used as proof of accuracy and priority.

X “A recent poll shows that 60% of voters support the policy.”

Although temporally recent, this is straightforward factual reporting rather than a persuasive appeal to recency.

X “In a recent speech, a respected senator criticized the plan and rightfully so.”

Although it mentions recency, the persuasive force comes from the senator’s authority and opinion rather than the timing. The temporal marker is incidental, not persuasive.

B.3 Urgency

Definition: *Urgency framing creates a sense of immediacy or imminence by emphasizes limited time, decisive moments, ticking clocks, last chances, ultimatums, or imminent threats.*

Rhetorical function: *It depicts situations as temporally constrained or approaching a critical moment in order to heighten perceived stakes and pressure, or call for immediate action.*

✓ “Congress must act within 72 hours to prevent a government shutdown..”

A specific deadline creates immediacy and pressure for political decision-making.

✓ “This election is the defining moment for our nation’s future..”

Frames the present moment as uniquely decisive and urgent.

X “Most voters say healthcare is an important issue..”

Conveys significance, but without compressed time or urgency.

X “Only a limited number of senators support the bill..”

Indicates scarcity of support, but not temporal scarcity; no deadline or ticking clock.

B.4 Temporal Anchoring

Definition: *Temporal anchoring frames the discussion through the lens of past events that serve as interpretive anchors.*

Rhetorical function: *It activates shared memory and emotional resonance, using familiar reference points to legitimize current perspectives or priorities, or to bias interpretation through invocation of certain events in the past.*

✓ “We live in a post-9/11 world where national security must come first, and the congress must take immediate action.”

The landmark date frames a lasting shift in values and policy priorities.

- ✓ **“This legislation represents the boldest reform since the New Deal.”**

Frames present significance by situating it relative to a landmark era.

- ✗ **“Security policies were tightened after 9/11.”**

Merely states historical sequence.

- ✗ **“We live in a world where national security must come first, and the congress must take immediate action.”**

It is the same example from earlier, but the omission of the temporal reference to 9/11. So while it is persuasive, it is not anchored in time.

B.5 Nostalgia

Definition: *Nostalgia framing invokes a cherished past as an ideal or standard for the present or future.*

Rhetorical function: *It activates a sense of identity and belonging, calls for restoration and frames it as progress, and softens resistance to a familiar past.*

- ✓ **“We must return to the prosperity of the postwar years.”**

Frames a remembered era of economic success as the standard for present policy.

- ✓ **“We need to revive the spirit of bipartisanship that once defined Congress.”**

Invokes a remembered political culture as guidance for present reform.

- ✗ **“The city is restoring historical facades downtown.”**

Reports an action but does not frame the past as a persuasive political ideal.

- ✗ **“Many voters say they value tradition.”**

Expresses a general liking for continuity, but without invoking a shared past as a model for today.

B.6 Temporal Contrast

Definition: *Temporal contrast framing emphasizes change by juxtaposing different time periods such as “then” versus “now.”*

Rhetorical function: *It dramatizes contrast to mark decline, progress, or reversal, thereby justifying action or interpretation.*

- ✓ **“Once a neglected district, now a thriving hub.”**

Direct before–after comparison makes temporal change salient.

- ✓ **“Twenty years ago tuition was manageable, today it traps students in debt.”**

Uses then vs. now contrast to highlight deterioration.

- ✗ **“Contrasting 2005 to 2010, the downtown population increased by 20%.”**

Without full context, this sentence is a factual report of a statistical change without rhetorical contrast.

- ✗ **“Urban elites and rural voters have conflicting values, and policymakers must bridge this divide.”**

Makes a rhetorical appeal based on geographical contrast, not a temporal one.

B.7 Continuity

Definition: *Continuity framing expresses persistence across time, encompassing stability, stagnation, momentum, or recurrence.*

Rhetorical function: *It legitimizes through longevity, reassures through endurance, or normalizes repetition by presenting ongoing trajectories or cycles as natural and inevitable. It can also convey discontent, dissatisfaction, or frustration with an unchanging situation.*

- ✓ **“For centuries this constitution has safeguarded our liberties.”**

Longevity is invoked as a direct source of legitimacy, even though whether the constitution truly safeguarded liberties is a matter of interpretation and debate.

- ✓ **“Our neutrality has endured for generations, and it continues to serve us well.”**

Persistence over time is framed as proof of value and stability.

- ✓ **“Every election season, the same promises return.”**

Uses recurrence as both explanation and expression of discontent and critique.

- ✓ **“Markets crash and recover; this downturn will pass.”**

Frames recurrence as reassurance.

- ✓ **“Our coalition has grown election after election, and nothing can stop this progress now.”**

Past trajectory is invoked to suggest inevitability.

- ✓ **“This campaign gains strength with every new supporter who joins us.”**

Accumulating participation is framed as persuasive momentum.

- ✗ **“The law remains on the books.”**

Factually reports longevity, but without using it in a persuasive way.

- ✗ **“The 1990 law is widely supported by voters.”**

Legitimacy is attributed to present popularity, not to persistence or longevity over time.

- ✗ **“As every winter, flu cases surge again.”**

Observes a seasonal pattern but does not employ it rhetorically.

- ✗ **“The flu is a serious threat to public health and cannot be ignored.”**

Makes an evaluative claim without invoking recurrence.

- ✗ **“The rally had great energy and everyone was fired up.”**

Shows force and enthusiasm in the moment, but not a trajectory over time.

- ✗ **“Voter turnout has increased in each of the last three elections.”**

Without the full context, this sentence merely reports momentum, but without any rhetorical appeals.

B.8 Skeptical

Definition: *Skeptical framing casts doubt about the future by emphasizing uncertainty, risk, pessimism, potential failure, or negative outcomes*

Rhetorical function: *It highlights vulnerability or doubt to provoke caution, hesitation, or reassessment of confidence in projected outcomes.*

- ✓ **“The plan may collapse under financial pressure..”**

Draws attention to risk in order to slow or block action.

- ✓ **“The merger could backfire on consumers..”**

Frames a potential outcome as harmful and uncertain.

- ✗ **“The plan costs 2 billion dollars..”**

A financial fact, not a speculative projection.

- ✗ **“As history shows, a ceasefire will eventually be signed..”**

Treats the outcome as inevitable certainty rather than a contingent risk.

C Additional Corpus Analyses

C.1 Unsupervised Document Tagging Reliability

As described in Section 4.1, we used LLM-assisted document-level opinion tagging to guide sampling, with the goal of enriching the dataset with articles likely to contain persuasive temporal framing. Prior to sampling, only 22.1% of English articles and 13.4% of German articles were tagged as opinionated. This indicated that such content formed a minority of the raw corpus. To ensure sufficient coverage, we upsampled documents tagged as opinionated while still retaining a substantial subset of non-opinionated ones to account for potential noise in these LLM generated tags.

After sampling, non-opinionated articles constituted 28.6% of the English subset and 25.5% of the German subset, resulting in final opinion-tag rates of 71.4% and 74.5%, respectively.

We now evaluate the reliability of this weak document-level signal by comparing LLM opinion tags against sentence-level human temporal framing annotations.

Using the presence of at least one temporally framed sentence as ground truth, 84.7% of documents tagged as opinionated by the LLM contained temporal framing, corresponding to a high true-positive rate for detecting framing-bearing articles.

In contrast, 41.9% of documents tagged as non-opinionated still exhibited temporal framing, indicating a substantial false-negative rate when opinionatedness is used as a proxy for framing presence. These results confirm that LLM-based opinion tagging is strongly correlated with temporal framing, but does not fully capture its distribution.

Beyond binary presence, documents tagged as opinionated also exhibited higher framing intensity. Across all documents, LLM-opinionated articles showed a higher mean ratio of framed sentences (0.150 vs. 0.067) and a higher mean number of temporal frame assignments per sentence (0.194 vs. 0.085) compared to LLM-non-opinionated articles. When restricting the comparison to documents containing at least one framed sentence, opinionated articles still displayed slightly higher framing density (mean framed-sentence ratio 0.177 vs. 0.161) and substantially more framed sentences per document on average (7.55 vs. 4.37).

These findings corroborate the soundness of our sampling strategy: while LLM opinion tags were effective for enriching the dataset with framing-heavy documents, retaining a controlled proportion of LLM-negative articles was necessary to capture temporal framing in ostensibly factual reporting. LLM predictions were used exclusively for document-level sampling decisions; all temporal framing labels and intensity measures were derived solely from human annotations.

C.2 Framewise Inter-Annotator Agreement at the Sentence-Level

Temporal Frame	α_{English}	α_{German}
Primacy	0.506	0.768
Recency	0.469	0.302
Urgency	0.442	0.593
Temporal Anchoring	0.590	0.434
Nostalgia	0.416	0.666
Temporal Contrast	0.538	0.640
Continuity	0.552	0.630
Skeptical	0.443	0.674

Table 5: **Sentence-level per-frame inter-annotator agreement.** It is measured with Krippendorff’s α (nominal), treating each temporal frame as a binary label (present vs. absent).

Table 5 reports sentence-level inter-annotator agreement scores computed separately for each temporal frame. Agreement is measured using Krippendorff’s α (nominal), treating each frame as

a binary label (present vs. absent). These frame-wise scores complement the aggregate agreement results reported earlier in Table 2 and illustrate variability in annotation difficulty across temporal frames.

C.3 Inter-Annotator Agreement at the Document-Level

To compute document-level labels, we take, for each annotator, the union of sentence-level frame assignments within an article. At the document level (Tables 6–7), agreement increases, which indicates that annotators agree more on overall framing than on sentence boundaries. Document-level detection achieves $\alpha = 0.679$ (EN) and $\alpha = 0.802$ (DE), while mean per-frame agreement increases to 0.579 (EN) and 0.692 (DE). Document-level set agreement remains high, with Set-F1₊ = 0.761 for English and 0.710 for German. Finally, document-level framing intensity in Table 8 shows rank agreement across annotators, with Spearman correlations of 0.575 and 0.541 (EN) and 0.839 and 0.805 (DE) for framed-sentence and frame-assignment counts, respectively.

Lang.	N_{docs}	α_{det}	$\bar{\alpha}_{\text{frame}}$	min	max	Set-F1 ₊
English	50	0.679	0.579	0.431	0.702	0.761
German	51	0.802	0.692	0.405	0.880	0.710

Table 6: **Document-level inter-annotator agreement.** Document-level labels are obtained by taking, for each annotator, the union of sentence-level temporal frame assignments within an article. Metrics are defined analogously to the sentence-level setting.

Temporal Frame	α_{English}	α_{German}
Primacy	0.620	0.880
Recency	0.467	0.405
Urgency	0.658	0.642
Temporal Anchoring	0.432	0.623
Nostalgia	0.679	0.652
Temporal Contrast	0.702	0.680
Continuity	0.638	0.795
Skeptical	0.431	0.863

Table 7: **Document-level per-frame inter-annotator agreement.** It is measured with Krippendorff’s α (nominal), using document-level frame presence derived from sentence-level annotations.

Lang.	ρ_{sent}	MAE _{sent}	ρ_{frame}	MAE _{frame}
English	0.575	5.67	0.541	7.93
German	0.839	2.43	0.805	3.78

Table 8: **Document-level agreement on temporal framing intensity.** ρ denotes Spearman rank correlation and MAE the mean absolute error. Subscripts indicate whether scores are computed over framed sentences or frame assignment counts.

C.4 Temporal Frame Co-occurrence

Figure 4 shows the sentence-level co-occurrence patterns of temporal frames in the corpus. Values are log-scaled to highlight relative associations between frame pairs and illustrate the multi-label nature of temporal framing.

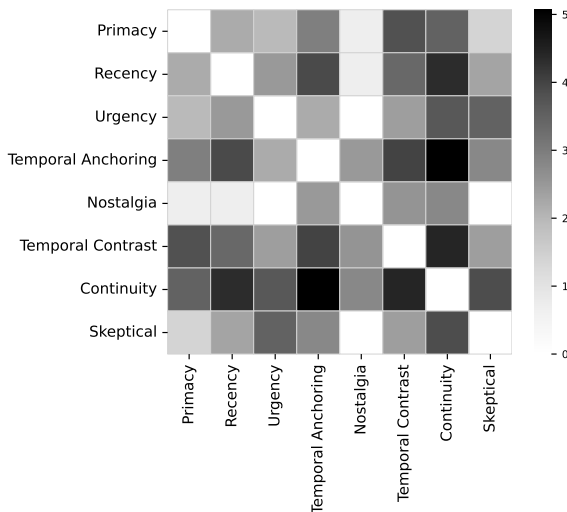


Figure 4: **Log-scaled heatmap of co-occurrence of temporal frames in the corpus.**

C.5 Corpus Splits

	Train	Dev	Test	All
English	138/166 (83.1%)	27/33 (81.8%)	33/39 (84.6%)	238
German	95/153 (62.1%)	18/30 (60.0%)	24/37 (64.9%)	220
Total	233/319 (73.0%)	45/63 (71.4%)	57/76 (75.0%)	458

Table 9: **Corpus distribution across train, dev, and test splits.** Entries show *framed/total* documents per split (percentage in parentheses); for example, *138/166 (83.1%)* indicates that 138 of 166 English training articles contain at least one temporal framing instance.

Table 9 reports the distribution of documents across the train, development, and test splits for each language, including the number and proportion of documents containing at least one tem-

porally framed sentence. The splits were constructed to preserve, within each language, the relative prevalence of temporally framed and non-framed documents, while maintaining an approximate 70/13/17 train-development-test ratio.

C.6 Lexical Analysis

Temporal Frame	Most Associated Terms
Primacy	<i>record, arrivals, first, ever, democratic, lot, never, history, ufc, high, year, since</i>
Recency	<i>recently, latest, recent, navy, southern, hours, weeks, week, tariffs, friday, brought, study</i>
Urgency	<i>clinicians, essential, medical, must, soon, threat, rise, citizens, critical, needs, military, today</i>
Temporal Anchoring	<i>neocon, battle, copy, century, civil, refused, term, harry, pandemic, rural, reminder, remained</i>
Nostalgia	<i>you, have</i>
Temporal Contrast	<i>previous, longer, amendment, ago, plastic, now, once, become, today, until, research</i>
Continuity	<i>steady, increasingly, repeatedly, streak, spirit, methods, scandal, remained, cryptocurrency, continues, remains, volatile</i>
Skeptical	<i>non-citizens, spouses, affect, non-citizen, qdot, estate, laws, spouse, senators, planning, broader, tariffs</i>

Table 10: **Top terms most strongly associated with each temporal frame, estimated using log-odds keyness.** Words are shown in descending order of association strength.

To better understand how temporal framing is realized at the surface level, we conduct a lexical analysis of sentences annotated with temporal frames. Rather than attempting to define frames purely in lexical terms, this analysis aims to identify words that are disproportionately represented in sentences annotated with a given frame relative to the rest of the corpus. We quantify these associations using smoothed log-odds ratios (Monroe et al., 2008), which assign higher scores to words that occur more frequently in sentences expressing a given frame than would be expected based on their overall corpus frequency.

The resulting lexical profiles in Table 10 summarize the most strongly associated words for each temporal frame and reveal that most frames ex-

hibit clear and intuitive surface realizations. For example, Recency is strongly associated with words such as *recently*, *today*, and *latest*, while Urgency is characterized by markers of necessity and imminence such as *must*, *critical*, and *soon*. Temporal Contrast similarly shows distinctive cues related to comparison across time, including *previous*, *once*, and *now*.

However, Temporal Anchoring, Skeptical, and Nostalgia frames exhibit comparatively weaker lexical signals in the log-odds analysis. For Temporal Anchoring, this outcome is expected given its primary function of situating claims with respect to specific events, periods, or historical reference points. As a result, anchoring expressions naturally co-occur with event- or domain-specific vocabulary, rather than drawing from a small, recurring set of surface cues.

For Skeptical framing, we complement log-odds with a lexicon-based diagnostic targeting hedging and uncertainty terms. Such cues occur in 49.2% of Skeptical-labeled sentences, compared to 10.8% elsewhere in the corpus (a 4.56× enrichment). This signal is driven primarily by modal verbs expressing contingency (e.g., *could*, *would*, *may*, *might*) as well as explicit uncertainty terms (e.g., *risk*, *concern*, *uncertainty*). This indicates that Skeptical framing is lexically realized through uncertainty and qualification rather than topical content.

Nostalgia framing in our corpus is primarily realized through evaluative and commemorative language rather than explicit temporal expressions. Nostalgia-labeled sentences frequently invoke idealized past states, historical figures, or perceived loss (e.g., references to legacy, vanished glory, or return), often without overt temporal adverbials such as *then* or *back*. Consequently, simple temporal cue-based analysis tends to under represent Nostalgia. Notably, the log-odds analysis for Nostalgia shows words such as *you* and *have*, which reflect the topical prevalence of commemorative and obituary-style writing within Nostalgia-labeled instances rather than nostalgia-specific lexical cues.

D Social Science Roots of the Temporal Framing Taxonomy

In this section, we ground each temporal frame from our taxonomy to foundational theories from psychology, cognitive science, communication, sociology, and political science. These theoretical underpinnings shed light on why certain temporal

frames resonate so strongly with audiences, based on how people process time-related information in memory, attention, and decision-making. We further illustrate how such concepts map onto the frame's persuasive function.

D.1 Primacy – First Impressions and Anchoring

The *Primacy* frame, which attributes special significance to being first in time, leverages one of the most established cognitive biases: the primacy effect. In psychology, the primacy effect refers to the tendency for information encountered earlier to be weighted more heavily than information encountered later (Asch, 1946). Early inputs receive greater attention and rehearsal, making them more likely to be encoded into long-term memory (Feigenbaum and Simon, 1962; Murdock Jr, 1962). Once an initial interpretive framework is established, later details are often assimilated into that framework rather than fundamentally revising it.

Classic work by Asch (1946) demonstrates that describing an individual with positive traits first leads to a more favorable overall impression than presenting the same traits in reverse order. In the news context, a Primacy frame can set the tone by emphasizing a first event, precedent, or milestone, thereby adjusting audience expectations for everything that follows. This connects closely to the anchoring heuristic in decision-making, where initial information serves as a mental benchmark that biases subsequent judgments (Tversky and Kahneman, 1974; Kahneman and Tversky, 1979; Furnham and Boo, 2011). By highlighting the first discovery, first warning, or unprecedented event, communicators tap into an audience's cognitive bias to favor the starting point in a sequence. Primacy framing draws on a robust cognitive tendency whereby what comes first enjoys a durable advantage in memory, interpretation, and influence.

D.2 Recency – Latest Events and the Availability Heuristic

The *Recency* frame emphasizes the most recent developments and is grounded in well-established findings from memory research (Feigenbaum and Simon, 1962; Murdock Jr, 1962). The recency effect describes the improved recall of information presented most recently, particularly when no substantial delay intervenes. Beyond recall, recent events exert disproportionate influence through

the availability heuristic (Tversky and Kahneman, 1973), whereby people estimate importance, likelihood, or frequency based on how easily examples come to mind.

Because recent or vivid incidents are more accessible in memory, they are often perceived as more important or representative than older information. This can distort risk perception. For example, a surge of recent crime reports may lead people to believe crime is increasing, even when long-term trends show stability or decline. The Recency frame capitalizes on this tendency by foregrounding today's figures or latest developments as the most relevant indicators of reality. In media discourse characterized by continuous news cycles, the newest headline often displaces earlier context, which reinforces significance through recency. Overall, the persuasive force of the Recency frame lies in our inclination to treat what just happened as especially informative, sometimes at the expense of broader temporal context.

D.3 Urgency – Imminence, Time Pressure, and Present Bias

The *Urgency* frame highlights limited time or imminent consequences and draws on both classical rhetoric and contemporary behavioral science. In ancient Greek rhetorical theory, the distinction between *chronos* and *kairos* provides a foundational account of how time functions persuasively (Smith, 1969). *Chronos* refers to chronological, sequential time measured in objective units such as dates, durations, or intervals. By contrast, *kairos* denotes the qualitatively “right” or opportune moment for action, which emphasizes timeliness, and consequence rather than metric placement on a timeline.

Kairotic appeals frame moments as fleeting, critical, or irrevocable, creating a sense of pressure that can be highly motivating. While references to *chronos* may appear on the surface, such as deadlines or countdowns, the persuasive force of urgency arises when these references are rhetorically transformed into kairotic claims about actionability. The Urgency frame thus operates primarily through *kairos* rather than *chronos*, by using time as a rhetorical device rather than a neutral dimension.

In psychology and marketing, this logic is echoed in scarcity principles (Cialdini, 2001). When opportunities or resources are framed as time-limited, they trigger fear of missing out and prompt rapid decision-making. From a behavioral

economics perspective, the effectiveness of Urgency is closely tied to present bias (Ainslie, 1975; O'Donoghue and Rabin, 2015), which describes the tendency to give more significance to immediate rewards or threats than future ones. People prefer smaller rewards now over larger rewards later, and they react more strongly to dangers portrayed as imminent than to those framed as distant (Trope and Liberman, 2010).

By compressing the temporal horizon, the Urgency frame counteracts temporal discounting, which makes future consequences feel concrete and immediate. By emphasizing that the next hours or days are critical focuses attention, it heightens the sense of alert, and encourages action. Whether framed through *kairos* or present bias, by making time scarce, one reliably intensifies engagement and motivates immediate response.

D.4 Temporal Anchoring – Contextualizing the Present through the Lens of the Past

Temporal Anchoring contextualizes present events by anchoring them to past moments. This strategy mirrors the anchoring effect identified by Tversky and Kahneman (1974), whereby judgments are disproportionately influenced by an initial reference point. In the temporal domain, a past event could serve as a powerful anchor that shapes interpretation of current developments.

Once an anchor is established, subsequent information is evaluated relative to it. Referring to a contemporary issue as occurring in a “post-9/11 world,” for instance, activates a rich schema associated with that historical event, and it guides interpretation and emotional response. Schema theory (Carbon and Albrecht, 2012) helps us understand that when one anchors to past events, such anchoring activate networks of associated beliefs that structure understanding of new information.

In political communication and sociology, historical analogies are widely used to legitimize or critique present actions (Wodak, 2015). By anchoring to significant moments in the past such as past crises, or previous conflicts, one transfers the schematic memory from that past to the present situation. Temporal anchors are often emotionally charged, drawing on wars, disasters, or triumphs to imbue current issues with significance. The persuasive power of this frame lies in our reliance on the past as a reference point for making sense of the present. By establishing history as an anchor, communicators guide audiences toward a preferred

interpretation of ongoing events.

D.5 Nostalgia – Rose-Tinted Memory and Positive Past Identity

The *Nostalgia* frame invokes an idealized past and resonates strongly due to the reconstructive nature of human memory (Starobinski, 1966; Mitchell et al., 1997). Nostalgia is not simple recollection of past events but an emotion-driven reconstruction in which negative details are often downplayed and positive aspects are amplified. This reconstructive bias produces a rosy, idealized image of earlier times, which makes them appear more appealing than the present.

Social psychological research shows that nostalgia serves important identity functions. It induces a sense of self-continuity across time by linking present identity to valued aspects of the past (Sedikides et al., 2008). By recalling where individuals or groups came from, nostalgia reinforces the idea of a stable core identity. A Nostalgia frame thus promises that returning to past values or conditions can restore comfort, meaning, or greatness.

Nostalgia also connects with the concept of collective memory (Halbwachs, 1950) where societies construct shared narratives of a golden age. These narratives foster social cohesion and increase optimism and resilience during periods of uncertainty. Empirical work suggests that nostalgia increases feelings of social connectedness and meaningfulness, functioning as an existential resource (Wildschut et al., 2006; Sedikides et al., 2004). In political communication, nostalgic rhetoric is especially prominent in populist narratives because it is emotionally potent and identity-affirming (Wodak, 2015; Taggart, 2000). By painting the past in positive hues, the Nostalgia frame activates longing for what is perceived as lost and motivates efforts to reclaim it.

D.6 Temporal Contrast – Then vs. Now Comparisons and Temporal Comparison Theory

The *Temporal Contrast* frame juxtaposes past and present, highlighting change over time. This approach relates to the Temporal Comparison Theory, which extends social comparison principles to comparisons across time (Albert, 1977; Suls and Wheeler, 2000). Just as individuals evaluate themselves relative to others, they also evaluate present circumstances relative to past or anticipated future states.

Temporal comparisons can be upward or downward. Upward temporal comparisons frame the past as better than the present, which elicits dissatisfaction, pessimism, or alarm. Downward temporal comparisons frame the past as worse, which promotes feelings of progress, relief, or hope (Wills, 1981). Statements such as “once a thriving hub, now a city in decline” exemplify upward temporal comparison.

These effects parallel findings in social comparison research, where upward comparisons can motivate or distress, and downward comparisons tend to bolster self-evaluation (Wilson and Ross, 2001). By placing two time points side by side, they invite comparative inferences about trajectory, decline, or progress. In this sense, Temporal Contrast frames draw persuasive force from a basic cognitive tendency toward reference-dependent evaluation where judgments are anchored to comparison points rather than being assessed in isolation (Kahneman and Tversky, 1979).

D.7 Continuity – Reassurance and Fatigue in Persistent Systems

The *Continuity* frame emphasizes persistence across time, encompassing stability, stagnation, momentum, or recurrence. It is realized by presenting conditions, processes, or trajectories as enduring, ongoing, or repeatedly recurring (e.g., “X has been true for decades,” “history is repeating itself,” or “this trend continues”). While continuity can function to reassure or legitimize through longevity, we observed that in our corpus it is most frequently used to highlight stagnation, failure to progress, or the persistence of undesirable states.

From a cognitive and decision-science perspective, continuity connects to status quo bias, whereby individuals systematically prefer existing arrangements and resist change, even when alternatives may be superior (Samuelson and Zeckhauser, 1988). This bias, reinforced by loss aversion (Kahneman and Tversky, 1979) and cognitive inertia (McGuire, 1960), helps explain why persistent conditions are often normalized and why continuity can be framed as inevitable rather than contestable. System justification theory further explains how endurance itself becomes a source of legitimacy, leading individuals to defend and rationalize longstanding social or institutional arrangements (Jost and Banaji, 1994).

Beyond stagnation, continuity also manifests through cyclical or recurrent temporal models, in

which repetition is framed as expected. When communicators invoke recurrence or cycles, continuity framing normalizes repetition and downplays the possibility of rupture or transformation. In political contexts, such framing can suggest that outcomes are unsurprising as they have already happened before, and are therefore historically patterned and hence are expected to continue to happen.

Continuity can also express momentum. In this case, continuity framing highlights trends that continue to rise or decline by embedding them within an unfolding time trajectory. Such framing implies that change is either already underway or constrained by prior momentum.

Overall, Continuity framing derives persuasive force by presenting time as unbroken, patterned, or accumulating. Whether framed as legitimate endurance, inevitable recurrence, forward momentum, or frustrating stasis, continuity invites audiences to interpret the present as tightly constrained by what has persisted before.

D.8 Skeptical – Doubting the Future, Risk Aversion, and Negativity Bias

The Skeptical frame casts doubt about the future (e.g., “XYZ plan may collapse under pressure”), and its impact is correlated to several psychological tendencies related to risk perception and negative information processing. One key principle at play is loss aversion, a cornerstone of prospect theory (Kahneman and Tversky, 1979). Loss aversion holds that people feel potential losses more acutely than equivalent gains. Therefore, a message focusing on what could go wrong or what one might lose if a scenario fails will often resonate strongly. By framing the future in terms of possible failure or pitfalls, the Skeptical frame leverages our predisposition to avoid loss and pain. Audiences are likely to give more weight to warnings of collapse than to equally plausible assurances of success because, psychologically, avoiding a bad outcome is more motivating than securing a good one. Closely related is the broader negativity bias in human cognition: negative information (warnings, dangers, criticisms) generally draws greater attention and has a bigger impact on judgments than positive information (Kanouse and L. Reid Hanson Jr., 1987). In communications, this is reflected in the news media’s bias by injecting doubt and highlighting uncertainty about the future. Essentially, it tells the audience: do not get too comfortable, problems likely lie ahead. This can prompt more critical eval-

uation and vigilance. Another theoretical link is ambiguity aversion where people are uneasy about unknown outcomes and often prefer known risks over unknown ones (Fox and Tversky, 1995). A Skeptical frame accentuates the unknown (“will the plan collapse? It’s uncertain”), potentially triggering a cautionary stance where people would rather err on the side of caution than bet on an overly optimistic plan. Skeptical framing derives its persuasive form by framing future developments as tenuous or doubtful, thus allowing communicators to justify conservative or preventive actions in the present.

E Additional Experimental Details

E.1 XLM-R Training Objective and Optimization Details

We next describe the optimization and training strategies used to address extreme label imbalance and sparsity in sentence-level temporal framing. In particular, we detail (i) a weighted Binary Cross-Entropy objective that corrects for frame-level imbalance across the corpus (Section E.1), and (ii) a label-aware batching scheme that mitigates batch-level sparsity during training (Section E.1). While loss reweighting adjusts the relative contribution of rare temporal frames, label-aware batching ensures consistent exposure to positive supervision at the mini-batch level. Together, these mechanisms stabilize training in a multi-label setting where positive instances are both rare and unevenly distributed across frames.

Weighted Binary Cross-Entropy Objective Let K denote the number of temporal frames. For each sentence, the model produces logits $\mathbf{z} \in \mathbb{R}^K$, which are converted to probabilities via independent sigmoid activations. Ground-truth labels are represented as $\mathbf{y} \in \{0, 1\}^K$, allowing multiple frames to be active simultaneously.

We optimize a weighted Binary Cross-Entropy loss defined as

$$\mathcal{L} = - \sum_{k=1}^K [w_k y_k \log \sigma(z_k) + (1 - y_k) \log(1 - \sigma(z_k))]$$

where $\sigma(\cdot)$ denotes the sigmoid function and w_k is a positive class weight for frame k .

The weight w_k is computed from the training data based on label prevalence:

$$w_k = \frac{N - n_k}{n_k},$$

where N is the total number of training instances and n_k is the number of sentences annotated with frame k . This weighting increases the contribution of rare temporal frames while leaving negative instances unweighted. The objective corresponds to standard Binary Cross-Entropy when $w_k = 1$.

Label-Aware Batching under Sparse Supervision

In addition to loss reweighting, training is affected by extreme label sparsity: under naive random batching, many mini-batches contain no positive labels, resulting in weak or uninformative gradient signals. To address this, we employ a label-aware batching strategy that enforces a fixed proportion of temporally framed sentences per batch.

Concretely, each mini-batch is constructed to contain a predefined fraction of sentences with at least one active temporal frame, with remaining batch slots filled by non-framed sentences. This strategy ensures consistent exposure to positive supervision while preserving the overall distribution of negative instances. Importantly, this batching procedure affects only training and does not alter the evaluation distribution.

In our experiments, we fix the proportion of temporally framed sentences to 50% per mini-batch. Once all available positive instances have been exhausted, no additional batches are constructed for that epoch. Negative instances are sampled randomly without replacement. As a result, each training epoch may not cover the full dataset, and the set of negative sentences observed can vary across epochs, introducing additional stochasticity while maintaining balanced supervision.

Hyperparameters and Randomness Control

All experiments use the default set of hyperparameters unless stated otherwise. Models are fine-tuned using the AdamW optimizer with a learning rate of 2×10^{-5} and a batch size of 16. Inputs are truncated to a maximum length of 512 tokens. Training proceeds for up to 20 epochs with early stopping based on development-set micro-averaged F1 score, using a patience of three epochs.

The optimizer is instantiated with default AdamW parameters: $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$, and no weight decay.

To address label sparsity during training, we employ a label-aware batching strategy with a fixed batch size of 16, constructed to include 50% sentences with at least one active temporal frame and 50% non-framed sentences. Positive and negative instances are sampled without replacement, and

Context Mode	Micro P	Micro R	Micro F1	Macro F1	Cost (\$)
Sentence-only	0.212	0.476	0.294	0.251	3.50
Prev + Target + Next	0.226	0.382	0.284	0.251	3.84
Full document	0.216	0.411	0.283	0.252	11.41

Table 11: **Effect of context granularity on zero-shot temporal framing classification.**

batch construction stops once all positive instances have been exhausted for a given epoch. As a result, individual epochs may not cover the entire training set, and the sampled negative instances may vary across epochs.

For reproducibility, we fix the random seed to 42 for all experiments. The seed is applied consistently across Python’s random module, NumPy, and PyTorch.

All reported results correspond to the model checkpoint achieving the best performance on the development set under this fixed hyperparameter and seed configuration.

E.2 LLM Experimental Details

E.2.1 Context Ablation

Based on initial explorations with larger context windows using GPT-5.2 on the development set, we feed models only the target sentence itself as input, denoted SENT (sentence-only), in all experiments. In this ablation, we compared SENT against two context-augmented variants: adding neighboring sentences (NEIGH) and providing the full article text (DOC). Adding context yielded no clear gains in overall detection performance and in some cases even reduced precision, likely due to irrelevant information. As illustrated in Table 11, we observed that surrounding context can make the model more conservative, leading it to miss frames that are already evident from the sentence alone. The sentence-only context achieves the highest performance while also being the most cost-effective. We, therefore, opt to use the SENT configuration for both fine-tuning and zero-shot evaluations.

E.2.2 LLM Inference Details

We use vLLM (Kwon et al., 2023) for efficient batch inference of all LLM-based models, including both zero-shot and fine-tuned variants. vLLM enables high-throughput generation through continuous batching and PagedAttention, allowing us to process data points efficiently. The random seed is fixed at 239 for reproducibility. The instruction set is shown in Appendix G.

Sampling Configuration. We use consistent sampling parameters across all experiments: temperature $\tau = 0.0$, top-p (nucleus sampling) $p = 0.8$, and top-k $k = 20$. The maximum generation length is set to 1024 tokens. Repetition, presence, and frequency penalties are disabled to avoid interfering with structured output generation.

Guided JSON Decoding. To ensure well-formed outputs, we employ vLLM’s structured output feature with a JSON schema derived from a Pydantic model. The schema constrains the model to output a JSON object with a single key `temporal_frames` containing a list of valid frame types. Valid frames are restricted to the eight categories: *Continuity*, *Nostalgia*, *Primacy*, *Recency*, *Skeptical*, *Temporal Anchoring*, *Temporal Contrast*, and *Urgency*. This guided decoding eliminates parsing errors from malformed JSON and prevents hallucinated frame categories.

E.2.3 LLM Fine-Tuning Details

We fine-tune Qwen3-8B and Llama-3.1-8B-Instruct using QLoRA (Dettmers et al., 2023), a parameter-efficient fine-tuning method that combines 4-bit quantization with low-rank adapters. Our implementation uses Unsloth¹ for efficient LoRA training.

LoRA Configuration. We apply low-rank adapters to all attention and feed-forward modules: `q_proj`, `k_proj`, `v_proj`, `o_proj`, `gate_proj`, `up_proj`, and `down_proj`. The LoRA hyperparameters are: rank $r = 16$, scaling factor $\alpha = 16$, and dropout 0.0. We use Unsloth’s gradient checkpointing for memory efficiency.

Training Configuration. Models are trained for 3 epochs with a per-device batch size of 16 and gradient accumulation over 4 steps, yielding an effective batch size of 64. We use the AdamW optimizer, a learning rate of 2×10^{-4} with linear decay, and 10 warmup steps. Weight decay is set to 0.01 and gradient clipping to 1.0. The random seed is fixed at 239 for reproducibility.

Response-Only Training. Following best practices for instruction tuning, we apply response-only training, which masks the loss on instruction tokens and computes gradients only on the assistant’s response.

¹<https://github.com/unslothai/unsloth>

Split	Model	Pri	Rec	Urg	Anch	Nos	Ctr	Cnt	Skp
EN + DE	<i>Zero-shot</i>								
	LLaMA-3.1-8B	0.11	0.20	0.06	0.13	0.10	0.14	0.16	0.00
	LLaMA-3.3-70B	0.16	0.17	0.24	0.21	0.19	0.23	0.33	0.17
	Qwen3-8B	0.03	0.10	0.16	0.15	0.11	0.23	0.27	0.07
	Qwen3-32B	0.19	0.28	0.18	0.26	0.16	0.23	0.31	0.18
	Qwen3-235B	0.16	0.22	0.21	0.24	0.32	0.26	0.30	0.12
	GPT-5.2	0.18	0.25	0.20	0.23	0.07	0.27	0.29	0.11
	<i>Supervised</i>								
	LLaMA-3.1-8B-Instruct	0.41	0.45	0.32	0.38	0.00	0.42	0.50	0.34
	Qwen3-8B	0.39	0.49	0.31	0.34	0.00	0.41	0.55	0.28
EN	<i>Zero-shot</i>								
	LLaMA-3.1-8B	0.16	0.28	0.12	0.25	0.19	0.20	0.22	0.00
	LLaMA-3.3-70B	0.20	0.25	0.30	0.28	0.30	0.42	0.40	0.17
	Qwen3-8B	0.08	0.21	0.25	0.25	0.26	0.30	0.35	0.05
	Qwen3-32B	0.28	0.44	0.25	0.34	0.19	0.29	0.32	0.10
	Qwen3-235B	0.26	0.36	0.29	0.37	0.38	0.36	0.36	0.05
	GPT-5.2	0.23	0.31	0.28	0.33	0.18	0.36	0.41	0.15
	<i>Supervised</i>								
	LLaMA-3.1-8B-Instruct	0.34	0.49	0.34	0.42	0.00	0.46	0.53	0.34
	Qwen3-8B	0.43	0.52	0.24	0.40	0.00	0.48	0.60	0.22
DE	<i>Zero-shot</i>								
	LLaMA-3.1-8B	0.06	0.06	0.02	0.04	0.06	0.00	0.05	0.00
	LLaMA-3.3-70B	0.09	0.06	0.14	0.14	0.13	0.08	0.20	0.18
	Qwen3-8B	0.01	0.03	0.08	0.07	0.04	0.13	0.16	0.08
	Qwen3-32B	0.10	0.09	0.09	0.17	0.12	0.10	0.28	0.29
	Qwen3-235B	0.06	0.07	0.13	0.13	0.27	0.16	0.18	0.23
	GPT-5.2	0.11	0.14	0.12	0.13	0.00	0.16	0.17	0.08
	<i>Supervised</i>								
	LLaMA-3.1-8B-Instruct	0.67	0.20	0.27	0.10	0.00	0.20	0.33	0.35
	Qwen3-8B	0.30	0.36	0.48	0.07	0.00	0.26	0.40	0.40

Table 12: **Per-label F1 scores across language evaluation splits.** Cnt = *Continuity*, Nos = *Nostalgia*, Pri = *Primacy*, Rec = *Recency*, Skp = *Skeptical*, Anch = *Temporal Anchoring*, Ctr = *Temporal Contrast*, Urg = *Urgency*.

Post-Training. After training, we merge the LoRA adapters into the base model weights using 16-bit precision (`merged_16bit`), producing a standalone model compatible with vLLM inference. The merged models are used directly with our zero-shot inference pipeline to ensure consistent evaluation.

E.3 Detailed Experimental Results

Following the results in Section 5.2, we provide additional experimental breakdowns by language and by temporal frame. Table 12 reports per-label F1 scores for all zero-shot and supervised models, broken down by evaluation split (English+German, English, and German). Table 13 shows binary detection and multilabel classification metrics under the same splits.

The per-label breakdown reveals substantial variation across temporal frames. Frames such as *Continuity*, *Temporal Anchoring*, and *Temporal Contrast* are detected more reliably across models and languages, particularly in English, where these frames occur more frequently and are associated with relatively stable lexical or structural cues. In contrast, low-frequency frames, most notably *Nostalgia*, exhibit consistently low F1 scores across both zero-shot and supervised settings, which highlights the strong effect of data sparsity and the dif-

Split	Model	Binary Detection			Multilabel Classification					
		P	R	F1	P	R	Micro-F1	Macro-F1		
EN + DE	Random Baseline	0.11	0.88	0.20	0.02	0.49	0.04	0.03		
	<i>Zero-shot models</i>									
		LLaMA-3.1-8B	0.14	0.94	0.24	0.06	0.34	0.10	0.11	
		LLaMA-3.3-70B	0.28	0.78	0.41	0.16	0.43	0.23	0.21	
		Qwen3-8B	0.21	0.75	0.33	0.08	0.29	0.13	0.14	
		Qwen3-32B	0.36	0.57	0.44	0.20	0.31	0.25	0.22	
		Qwen3-235B	0.31	0.70	0.44	0.17	0.38	0.24	0.23	
		GPT-5.2	0.34	0.67	0.45	0.24	0.46	0.31	0.29	
	<i>Supervised fine-tuning</i>									
		XLm-R	0.38	0.75	0.51	0.28	0.57	0.37	0.33	
		Llama-3.1-8B-Instruct	0.67	0.45	0.54	0.48	0.37	0.42	0.35	
		Qwen3-8B	0.60	0.54	0.57	0.48	0.40	0.44	0.35	
	EN	Random Baseline	0.20	0.88	0.32	0.03	0.48	0.06	0.06	
		<i>Zero-shot models</i>								
			LLaMA-3.1-8B	0.25	0.92	0.39	0.12	0.37	0.19	0.18
		LLaMA-3.3-70B	0.43	0.73	0.54	0.26	0.43	0.32	0.29	
		Qwen3-8B	0.41	0.69	0.52	0.18	0.29	0.22	0.22	
		Qwen3-32B	0.49	0.51	0.50	0.32	0.29	0.31	0.28	
		Qwen3-235B	0.49	0.64	0.56	0.30	0.36	0.33	0.30	
		GPT-5.2	0.46	0.58	0.52	0.37	0.44	0.40	0.26	
<i>Supervised fine-tuning</i>										
		LLaMA-3.1-8B	0.68	0.52	0.59	0.48	0.42	0.45	0.36	
		Qwen3-8B	0.67	0.58	0.62	0.54	0.42	0.47	0.36	
DE		Random Baseline	0.05	0.90	0.10	0.01	0.53	0.02	0.02	
		<i>Zero-shot models</i>								
			LLaMA-3.1-8B	0.07	0.99	0.12	0.02	0.25	0.03	0.04
			LLaMA-3.3-70B	0.16	0.91	0.27	0.07	0.45	0.12	0.13
		Qwen3-8B	0.11	0.93	0.19	0.03	0.30	0.05	0.07	
		Qwen3-32B	0.23	0.72	0.35	0.10	0.37	0.16	0.15	
		Qwen3-235B	0.18	0.85	0.29	0.08	0.43	0.13	0.15	
		GPT-5.2	0.21	0.83	0.33	0.14	0.64	0.23	0.17	
	<i>Supervised fine-tuning</i>									
		LLaMA-3.1-8B	0.61	0.24	0.34	0.46	0.20	0.28	0.26	
		Qwen3-8B	0.43	0.43	0.43	0.33	0.33	0.33	0.28	

Table 13: Overall results broken down by language evaluation splits.

ficuity of learning frames that rely on implicit evaluative references to the past.

Comparing results across languages highlights the impact of corpus composition. Across models, English consistently outperforms German, especially for recall and rare frames, where the German subset contains far fewer positive instances. Consequently, lower German performance reflects data sparsity rather than a failure of cross-lingual modeling.

Finally, the contrast between zero-shot and supervised models shows that task-specific training yields more balanced and consistent performance, though substantial per-frame variability remains, which reflects the discourse-level nature of temporal framing.

F Full Annotated Example

Violence Has Been Normalized

During the misnamed and mostly preposterous debate between Kamala Harris and Donald Trump, a moderator fact-checked Trump's claim that crime is up. In contrast to Trump's claim, moderator David Muir said that the FBI reports that crime is down, a claim that likely struck every viewer as obviously wrong. Shoplifting was not a way of life before lockdowns. **Nostalgia** **Temporal Anchoring** **Temporal Contrast** Most cities were not demographic minefields of danger around every corner. **Nostalgia** **Temporal Contrast** There was no such thing as a drugstore with nearly all products behind locked Plexiglas. **Nostalgia** **Temporal Contrast**

We weren't warned of spots in cities, even medium-sized ones, where carjacking was a real risk. **Nostalgia** **Temporal Contrast**

It is wildly obvious that high crime in the U.S. is endemic, with ever less respect for person and property. **Continuity** **Primacy** As for the FBI's statistics, they're worth about as much as most data coming from federal agencies these days. **Temporal Contrast**

They're there for purposes of propaganda, manipulated to present the most favorable picture possible to help the regime. Lies, Damn Lies and Government Statistics.

This is certainly true of the Bureau of Labor Statistics and the Commerce Department, which have been shoveling out obvious nonsense for years. **Continuity** Professionals in the field know it but go along for reasons of professional survival. In truth, we've never had a real economic recovery since lockdowns. **Continuity** **Temporal Anchoring**

Crime is up. **Temporal Contrast** Literacy is down. **Temporal Contrast** Trust has collapsed. **Temporal Contrast** Societies were shattered and remain so. **Continuity** **Temporal Contrast**

Only a few weeks following the officious fact-check at the debate, we now have new data from the National Crime Victimization Survey. **Recency** The Wall Street Journal reports: "The urban violent-crime rate increased 40% from 2019–2023. Excluding simple assault, the urban violent-crime rate rose 54% over that span. From 2022–2023, the urban violent-crime rate didn't change to a statistically significant degree, so these higher crime rates appear to be the new norm in America's cities."

But the FBI tries to tell you that crime is down. Sure, whatever they say.

The report isolates the "post-George Floyd protests" because no media source wants to mention the lockdowns. **Temporal Anchoring** It is still a taboo subject. **Continuity**

We somehow cannot say, even now, that the worst abuses of rights in U.S. history in terms of scale and depth were a disaster, simply because saying so implicates the whole of the media, both parties, all government agencies, academia and all the upper reaches of the social and political order.

Politics Has Become Life and Death

The problem of political division is getting alarmingly serious. **Continuity** **Urgency** It's no longer just about competing yard signs and loud rallies. **Temporal Contrast** We now have regular assassination attempts, plus even an extremely strange appearance of a bounty put on a candidate's head by an official agency. **Temporal Contrast**

Surveys have shown that 26 million people in the U.S. believe that violence is fine to keep Trump from regaining the presidency. **Continuity** Where might people have gotten that idea?

Probably from many Hollywood movies that fantasize about having killed Hitler before he accomplished his evil plus the nonstop likening of Trump to Hitler, and hence one follows from another. Liken Trump to Hitler and that is the result you produce.

There's private violence, public violence and many forms in between including vigilante violence. Rights violations against person and property are now normalized. **Temporal Contrast** This springs from the culture of our times which has been heavily informed and even defined by the deployment of state violence in service of policy goals, at a scale, scope and depth never before seen. **Continuity** **Primacy** **Temporal Contrast**

The Role of Censorship

Censorship is a major part of it. Censorship is the deployment of force in service of state power, and other institutions connected to state power, for purposes of culture planning. It's exercised by the shallow state, in response to the middle state, and on behalf of the deep state. It's a form of violence that interrupts the free flow of information: the ability to speak, and the ability to learn. Censorship trains the population to be quiet, afraid and constantly stressed, and it sorts people by the compliant versus the dissidents. Censorship is designed to shape the public mind toward the end of shoring up regime stability. Once it starts, there's no limit to it. **Continuity**

I've mentioned to people that Substack, Rumble and X could be banned by the spring of next year, and people respond with incredulity. **Skeptical** Why? Four years ago, we were locked in our homes and locked out of churches, and the schools for which people pay all year were shut down by government force. **Temporal Anchoring** If they can do that, they can do anything.

Remember Free Speech? **Nostalgia**

Censorship has been so effective that it's changed the way we engage with each other even in private. **Continuity** **Temporal Contrast**

Brownstone Institute, which I founded, recently held a private retreat for scholars, fellows and special guests. One very special guest wrote me that she was completely shocked at the freedom of thought and speech that was present in the room. As a mover in the highest circles, she had forgotten what that was like. This censorship coincides with a strange valorization of violence that we are presented with from all over the world: Ukraine, the Middle East, London, Paris and many American cities. Never have so many held video cameras in their pockets and never have there been so many platforms on which to post the results. **Primacy** **Temporal Contrast** One does wonder how all these relentless presentations of destruction and killing affect public culture. **Continuity**

Why They're Doing It

What purpose are all these soft, hard, public and private exercises of violence serving? The standard of living is suffering, lives are shortening, despair and ill health are main features of the population and illiteracy has swept through an entire generation. **Continuity**

The decision to deploy violence to master the microbial kingdom did not turn out well. Worse, it unleashed violence as a way of life. "When plunder becomes a way of life for a group of men in a society," wrote Frederic Bastiat, "over the course of time they create for themselves a legal system that authorizes it and a moral code that glorifies it."

That is precisely where we are. It's time we talk about it and name the culprit. **Urgency** Liberty, privacy and property were already unsafe before 2020 but it was the lockdowns that unleashed Pandora's box of evils. **Continuity** **Temporal Anchoring** **Temporal Contrast**

We cannot live this way. The only arguments worth having are those that name the reason for the suffering and offer a viable path back to civilized living. **Nostalgia**

Figure 5: A full annotated example illustrating temporal framing. Adapted from "Violence Has Been Normalized" by Jeffrey Tucker, originally published on *Daily Reckoning* (Oct. 11, 2024). URL: <https://dailyreckoning.com/violence-has-been-normalized/>.

G Prompt for Zero-Shot Experiments

Multi-label Temporal Framing Classification

You are an expert annotator for temporal framing in news text. Temporal framing refers to the rhetorical use of time-related elements to persuade, not merely to report chronology. A sentence should be labeled only if the temporal element contributes to its persuasive force.

Temporal framing taxonomy (short definitions and examples):
{BRIEF_TAXONOMY}

Key rules for annotation:

- Do not annotate factual chronological reporting or purely descriptive temporal mentions.
- Do not annotate quoted or indirect speech.
- If removing the temporal expression does not change the persuasive force, do not annotate.
- Multiple temporal frames may apply to a single sentence.

Context (optional):
{CONTEXT}

Target sentence:
{TARGET_SENTENCE}

Return the applicable temporal frames in structured JSON.

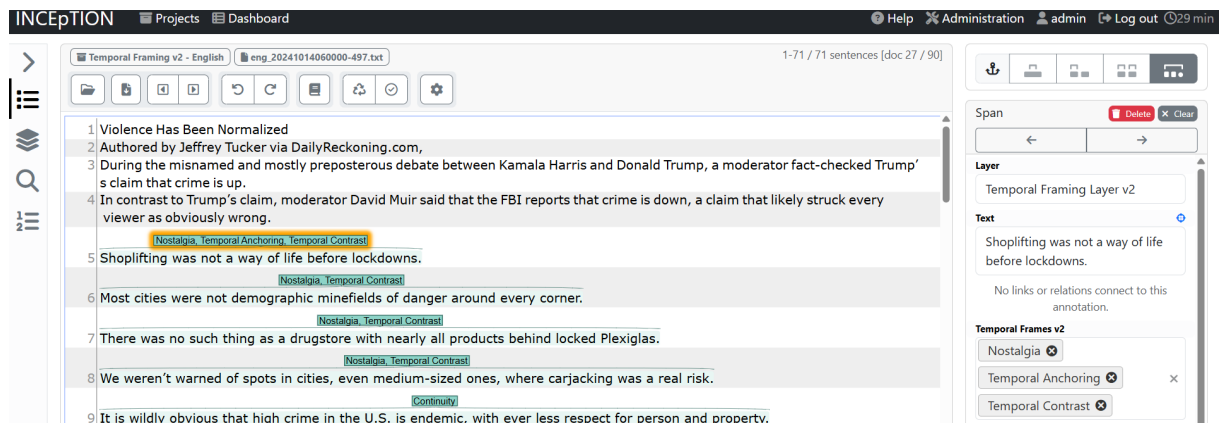
Figure 6: **Prompt template for temporal framing classification.** Taxonomy definitions from Figure 1b are inserted into the prompt via the BRIEF_TAXONOMY variable, after which the model predicts all applicable temporal frames in a single inference pass. The optional context block is included only in context-granularity ablation experiments.

H INCEpTION Annotation Platform

All annotations in this work were produced using the INCEpTION annotation platform, a web-based system designed for scalable, multi-layer linguistic annotation. INCEpTION supports span-based and sentence-level labeling, and configurable tagsets. These features make it suitable for complex annotation tasks that require consistency and quality control across multiple annotators and languages.

For temporal framing, we configured INCEpTION to operate at the sentence level, where annotators assign one or more temporal frame labels to a sentence. The interface allows annotators to view the full article context while focusing annotation decisions on individual sentences. The interface further facilitates workload assignment and monitoring, as well as inspection, revision, and adjudication of annotations.

Figure 7 shows the annotation interface as used in this work.



The screenshot displays the INCEpTION web interface. At the top, there's a navigation bar with 'INCEpTION', 'Projects', 'Dashboard', 'Help', 'Administration', 'admin', 'Log out', and '0:29 min'. Below this, a toolbar contains various icons for document management. The main area shows a list of 9 sentences from a document. Sentence 4 is highlighted, and several temporal frame labels are applied to it: 'Nostalgia', 'Temporal Anchoring', and 'Temporal Contrast'. On the right side, a 'Span' and 'Layer' panel is visible, showing the current layer as 'Temporal Framing Layer v2' and the text of the selected sentence. Below that, a 'Temporal Frames v2' panel lists the applied labels: 'Nostalgia', 'Temporal Anchoring', and 'Temporal Contrast', each with a close button.

Figure 7: **INCEpTION annotation interface.** The figure shows how the annotated example shown in Figure 1a looks like on the INCEpTION interface with sentence-level annotation and multiple temporal frame labels.