TRANSLINGUA, 2025

Vol 1, Issue 1, Pages 1-23

A Comparative Discourse Analysis of AI-Generated vs. Human-Produced News Reports: Evaluating Objectivity, Framing, and Ideology

Author/s: Amjad Naveed[1], Insha Ullah[2]

Affiliation: Department of English and Applied Linguistics, University of Lakki Marwat (ULM), Email: [1]amjadnaveed4star@gmail.com, [2]ullahinsha262@gmail.com

ABSTRACT: As large language models enter newsrooms, this study compares AI-generated and human-produced news on linguistic construction, framing, and objectivity. We analyzed N = 360 English-language reports (AI = 180; Human = 180) across politics, economy, science/technology, and health in the US, UK, and Pakistan (Jan–Jun 2025), combining CDA/SFL-based coding with computational indicators (lexical diversity/MTLD, syntactic complexity/MLC, readability, subjectivity/sentiment; quotation density, source diversity, reporting verbs). Tests include Welch's t with Benjamini–Hochberg correction, mixed-effects models (Origin fixed; Outlet random; Beat/Country controls), binomial GLMs for frames, an Objectivity Index, and an exploratory classifier. AI texts use more modality/hedging and slightly more passive voice but lower lexical diversity and syntactic complexity. The largest gaps are in sourcing (fewer quotes, sources, reporting verbs), yielding a lower Objectivity Index for AI. Human-Interest framing is more common in human stories, while AI tilts toward Economic Consequences. A classifier distinguishes origin with AUC ≈ 0.96. AI and human news form distinct discourse profiles; pairing AI drafting with sourcing protocols, disclosure, and frame-aware editing is recommended to meet journalistic standards.

Keywords: Framing, Ideology, Objectivity, Comparative Discourse Analysis, News Reports

To cite: Naveed, A. & Ullah, I (2025). A Comparative Discourse Analysis of AI-Generated vs. Human-Produced News Reports: Evaluating Objectivity, Framing, and Ideology. *TRANSLINGUA, 1*(1), pages 1-23.

1. INTRODUCTION

The emergence of generative artificial intelligence (AI) into news-making has intensified ancient discussions on how news language produces social reality and how much notions of objectivity, balance, and ideological neutrality are being operationalized or perverted in practice. Critical Discourse Analysis (CDA) conceives of news not as a transparent mirror of events but as a patterned form of public discourse, wherein linguistic choices (lexis, syntax, modality, transitivity) contribute to the framing of events and the positioning of audiences (Fairclough, 1995; van Dijk, 1988). In this school of thought, and based on Systemic Functional Linguistics (SFL), grammar of news is approached. Transitivity patterns that define actors and their actions, modality, which indicates commitment and evaluative attitude, and choices of lexicon, which define authorial position are used to operationalise meaning making in discourse (Halliday and Matthiessen, 2014; Fowler, 1991). Parallel framing theory explains the ways in which texts construct issues, diagnose, make moral judgments and mobilize remedy proposals (Entman, 1993; Gamson and Modigliani, 1989). Altogether, these theoretical approaches provide a strong set of prisms through which the similarities and differences between news created by AI and humans are interrogated.

The automation in journalism has been applied traditionally, but the range and the complexity of automation have grown significantly. Early robot journalism systems took structured data (sports statistics and corporate profits) and translated it into templated writing on a scale (Graefe, 2016; Dörr, 2016). Practitioner interviews and case studies suggest that automation can increase the speed and depth of coverage, but at the same time is generating transparency, journalistic judgment, and ethical responsibility problems (Diakopoulos, 2019; Thurman et al., 2017). The pre-LLM experimental and perceptual research-based evidence showed mixed responses of the audience; in certain situations, readers found the automated stories as informative and readable as human-created stories, although less exciting (Clerwall, 2014). A recently released meta-analysis has found that perceived quality depends on topical and contextual considerations, but the general acceptance of automated journalism by the audience is high (Graefe and Bohlken, 2020). However, with the recent development of large language models (LLMs) it has been possible to generate fluent, context-specific prose that goes far beyond the paradigms of data-to-text and makes a modern comparative discourse analysis not only timely but also urgent.

Located in the framework of Critical Discourse Analysis (CDA) and Systemic Functional Linguistics (SFL), the main question is how AI and human texts are different in the process of the linguistic construction of events. The early studies of news language show that lexical options, appraisal processes and clause structures are systematic encoding of ideological stances such as foregrounding this actor, backgrounding others, modulating claim, and normalising particular interpretation (Bell, 1991; Fowler, 1991; Halliday and Matthiessen, 2014; van Dijk, 1988). Template / prompt-based generation in automated texts can strengthen particular transitivity patterns, like tendency to agentless passives or nominalisations in general, and modal strategies typified by scrupulous hedges or default indicative. Big pretrained models, in contrast, can import the training learned stylistic priors (Recasens et al., 2013). The analysis of biased language through computational approaches has shown patterns

of subjective lexis, framing nouns, and markers of epistemic that can be measured quantitatively across corpora (Recasens et al., 2013; Wilson et al., 2005).

This comparative endeavour is further inspired by the normative ideal of objectivity. Objectivity historically became a professional standard during the transformation of the political economy and the transformations of the newsroom practices (Schudson, 1978), and is still a controversial but powerful standard in the journalism ethics (Maras, 2013). Objectivity may be operationalised by means of linguistic proxies, i.e. lower levels of attitudinal lexis, restrained modality, attribution of sources with precision and the balanced presence of actors and frames, which makes it subject to comparative analysis of AI-generated and human-written narratives. Empirical evidence is undergoing a development. The readability and perceived informativeness of some reporting areas (e.g., sports, finance) Pre-generative LLM Before access to generative LLMs, studies in newsrooms and readership implied that automation might be as effective as human performance on readability and perceived informativeness in some domains (e.g., sports, finance), but would pose concerns about nuance and news sense (Clerwall, 2014; Thurman et al., 2017; Graefe and Bohlken, 2020). The modern period of LLC showed that AI-generated articles can be challenging to discern as opposed to those written by humans and can be considered as equally credible, which is why it is time to investigate linguistic peculiarities, framing models, and ideological overtones (Kreps et al., 2022). According to recent audience research, there also exist settings that the output of the generative models is rated as more readable and with more textual structure than the human versions, even though the topic and outlet selection also play a role (Baptista et al., 2025). The only aspect that is not yet studied in depth is a comparative, systematic discourse analysis that goes beyond perceptions and investigates the ways that AI and human news are different in lexical choices, modality, and transitivity; the way that the differences between the two are correlated with the mechanisms of framing and ideological positioning; and how objectivity is constructed linguistically in both discourses.

This gap is filled in this research by using comparative discourse analysis on AI-generated and human-produced news reports. It draws upon CDA and SFL (Fairclough, 1995; Halliday and Matthiessen, 2014; van Dijk, 1988) and grounded in the framing theory (Entman, 1993; Gamson and Modigliani, 1989) by analyzing (1) linguistic construction (lexis, modality and transitivity); (2) framing and ideology (defining a problem, causal attributions, Through the triangulation of manual CDA coding with the computational measures (e.g., subjectivity lexicons and bias cues), the study will describe the systematic variation between AI and human news discourse instead of determining a hierarchical value on either.

1.2. Research Questions

i. Linguistic Construction of News: How do AI-generated news reports and human-generated news reports vary in the choice of lexicon, modality, and transitivity in the construction of news events?
ii. Framing and Ideology: What are the differences in the event framing and ideological positions of AI-generated and human-created news discourses?
iii. Objectivity and Bias: How objective are AI-generated news reports compared to human-generated reports, and how each type of discourse is biased?

2. LITERATURE REVIEW

This section of the paper presents a comprehensive description and analysis of the theory of the knowledge area and the empirical studies within the domain of knowledge.

2.1. News as discourse-analytic

An analytic perspective on the ways linguistic decisions create social reality would be suitable to a comparative inquiry of AI-generated versus human-generated news. Critical Discourse Studies (CDS) theorises news discourse as a location of co-construction of language, power and ideology and provides powerful procedures, including the discourse-historical approach, to bridge textual forms to socio-cultural situations (Wodak & Meyer, 2016). In CDS, corpus-assisted discourse studies (CADS) combine qualitative interpretation with corpus-based methods (keywords, collocation, concordance) to bring systematic regularities to the surface that would otherwise be difficult to notice in manual reading; it has shown itself to be scalable to large news corpora (Baker et al., 2008). In terms of news-specific discourse work, Richardson (2007) describes the encoding of stance by headlines, sourcing, and attribution, and Bednarek and Caple (2012, 2017) construct frameworks of analysis of news values (e.g., negativity, proximity) in both text and image, thus explaining how discourse constructs the newsworthiness. Together, CDS/CADS and news-discourse frameworks justify the study of lexical and sourcing and clause-level grammar of comparing AI and human news texts.

2.2. Linguistic means associated with (un)objectivity: evaluation, evidentiality, modality.

To operationalise the notion of objectivity, the evaluative language and source-of-knowledge marking must be taken into account. Studies of evaluation in news document how lexis and phraseology encode stance, gradability, and attitudinal positioning beyond overt opinion, often via lexical patterns and appraisal-like resources (Bednarek, 2006). Evidentiality and epistemic positioning Studies of English news language show that journalists indicate knowledge bases (e.g., attributed speech, documents, inference) and levels of commitment, thus influencing the perceptions of factuality and balance formed by the readers (Whitt, 2006). Computationally, factuality and commitment have been modeled using event-level annotations (e.g., FactBank) and certainty/polarity typologies, to create measurable proxies of text analytics objectivity (Sauri & Pustejovsky, 2009, 2012). The measurement of lexical evaluation, hedges/modals, and evidential attributions are encouraged by these traditions in comparing articles written by AI to those written by humans.

2.3. News and framing ideology: Concepts and computational measures.

In addition to the stance on the sentence level, framing theorises the processes through which texts emphasize specific definitions of the problems, their causes and moral judgments and solutions. Framing is consolidated as a pattern of media production and an effect on the audience by Scheufele (1999), but a content-analytic model (problem definition, causal interpretation, moral evaluation, treatment recommendation) that can be systematically coded is proposed by Matthes and Kohring (2008). Constructionist descriptions focus on culturally similar packages of frames (van Gorp, 2007). Baumer et al. (2015) compare approaches to

detecting framing language in political news in computational framing, and Card et al. (2015) make available the Media Frames Corpus to study cross-issue frames; Hamborg et al. (2021) provide materials to match news entities to framing clues. Neural architectures have been shown to be able to detect left/right slant using lexical-syntactic cues in text (Iyyer et al., 2014) and topic-specific sentiment distributions in haddad to detect ideological leanings (Bhatia et al., 2018). These streams offer confirmed tools frame size and lexico-syntactic cues, topic-sentiment curves to compare AI and human discourse on ideology and framing.

2.4. AI generated or automated journalism: production, perceptions, and roles.

Before the recent generation of large-scale language models, the roles and professional limits of automated journalism were charted through scholarship. The so-called robotic reporter is described by Carlson (2015) as a boundary object reorganising journalistic power and labour, and van Dalen (2012) discusses the re-definition of the core skills by the machine-written news and the resulting professional negotiation. The research on reader-perception has discovered that the perception of bias, credibility, and engagement shifts with the labeling of AI-written (or machine-written) content, where certain audiences perceive machines to be free of intent and, thus, it is not perceived as slanted (Wolker and Powell, 2018; Lee et al., 2017). These themes reemerge in recent commentaries on the generative-AI era, where journalists are reasserting expertise by way of verification, interpretation and transparency as large-language models creep on drafting work (van Dalen, 2024). Collectively, this literature makes AI a production support and a discursive subject whose appearance influences perceived objectivity and bias.

2.5. Language differences in AI- and human text.

Regardless of the news domain, systematic linguistic differences between AI- and human-written texts are consistently reported by independent multidimensional analyses. Based on Biber-style dimensions, Sardinha (2024) documents a misalignment in distribution of involved/informational, narrative, and online elaboration features, indicating that GPT-style outputs do not cluster as much as human registers. Massive comparisons also observe less stylistic diversity, narrower lexical distributions, and different sentence-level distributions in the output of LLM compared to human text (e.g., analyses of ACL SRW; larger-scale multi-domain studies), which also are relevant to the question of robotic uniformity (Zhang et al., 2023; Rocha and Mendes, 2025). In news in particular, syntactic/psychometric comparisons of LLM-generated versus human news yield significant differences, including in the use of personal pronouns, hedging, intensifiers, and sociolinguistic cues (e.g., personal pronouns, hedging, intensifiers), which allow reliable differentiation (Munoz-Ortiz et al., 2023; Zamaraeva et al., 2025). These results have informed our choice of lexical diversity, syntactic complexity and (epi)modal features as discriminators in comparative discourse analysis.

2.6. Bias, objectivity, and detection in Computational Linguistics

Computational research provides methodological instruments that are concordant with issues concerning objectivity that have historically been considered in journalism. There is empirical evidence that neural language models are more effective than lexicon-based baselines in identifying subtle, context-dependent bias cues (Hube and Fetahu, 2019; 2018). Moreover,

article text and engagement patterns can be used to predict media-level bias and factuality, thus, allowing corpus-scale estimates of editorial slant (Baly et al., 2018). To express event factuality, assertion strength, which are main constituents of what we call objectivity, the FactBank corpus offers proposition-level annotations that enable an analyst to measure commitment and polarity (Sauri and Pustejovsky, 2009). When comparing artificial intelligence and human content in the perception research, readers tend to show a bias towards the content by labeling it as human, even when hidden when making quality judgments, which also shows the presence of label-based bias in content judgments (Raman et al., 2025). When measuring objectivity and bias of AI-generated news over human news, the combination of these questioning areas enlightens both measurement (the required exact counts) and interpretation (audience reception) of results.

2.7. Transformation in framing and in LLM-Era.

With the growing use of large language models (LLMs) by newsrooms to rewrite and generate headlines, early studies show that affective reframing by AI can change the audience reaction-such as negativity-oriented reframing has been shown to boost click-through rates in recommender systems (Trattner et al., 2024). Experiments under controlled conditions of headlines demonstrate that there is a difference in perceived trust and effectiveness of AI-generated and human-crafted micro-texts (Spinde et al., 2025). On the audience level, preregistered research indicates that perceived quality does not affect willingness to read AI-generated news as much as it does disclosure and already held beliefs about AI, thus making it difficult to make simplistic claims like AI is less objective (Gilardi et al., 2025). These results highlight the importance of comparative discourse analysis to consider the textual characteristics and frames, as well as labeling strategies and contextual use of news to assess the objectivity, framing and ideology in AI- versus human-generated news.

3. METHODOLOGY

3.1. Research Design

Our design is a comparative, corpus-based mixed-method design that combines quantitative text analytics with the qualitative critical discourse analysis (CDA). The unit of measurement is the single straight-news piece (without including editorials and op-eds), which is in line with the accepted levels of transparency and reliability of content-analytic data (Krippendorff, 2019). To interpret qualitatively, we refer to the Systemic Functional Linguistics to analyze the types of transitivity / processes and participant roles and to the Appraisal framework to question the stance and evaluative language (Thompson, 2014; Martin and White, 2005). The concept of framing is operationalised by a commonly available collection of generic frames: attribution of responsibility, conflict, human interest, economic consequences, and morality, which is deductively applied and inductively extended by topic-specific sub-frames as the coder is trained (Semetko and Valkenburg, 2000). Triangulation of manual CDA with computational measures of subjectivity, modality/hedging, readability, lexical diversity and syntactic complexity is used to tackle the challenge of objectivity and bias (Hyland, 1998; Biber, 1995).

3.2. Corpus and Sampling

We assembled two matched corpora of January-June 2025. The human-generated corpus (HPC) is a collection of 180 straight-news stories sampled in six mainstream English-language news outlets in the United States, United Kingdom, and Pakistan (two per country), thus varying editorial conventions. The AI-generated corpus (AIC) consists of 180 reports based on (A) articles specifically marked by publishers as AI/AI-assisted, and (B) a controlled set of AI-generation AI-prompts and parameters, prompted to write same-day/same-topic stories in a fixed neutral news style; this dataset is archived to ensure reproducible results. To reduce topic effects, articles were stratified by beat (politics, economy/business, science/technology, health; 45 per beat [?]23 AI /22 human). Inclusion criteria: use of the English language, the approximate number of words was 300, the strait news format, not duplicates. Exclusion criteria included opinion/ analysis columns, editorials and live blogs. In each outlet and beat we random-sampled weekly, with day-of-week dispersion. In the case of AI-tagged supply shortages, the controlled set gave 1:1 topic matches to HPC. The last sample had N = 360 (180 in each of the classes), evenly balanced by beat, outlet, and country.

3.3. Measures

3.3.1. Manual CDA

We coded transitivity and agency (type of process; Actor/Goal/Sayer/Sensor; voice), Appraisal/stance (attitude, graduation, engagement), generic frames (and inductive sub-frames), and ideology cues (evaluation of actors/policies; choice of source/attribution patterns; antagonistic labels) (Thompson, 2014; Martin and White, 2005; Semetko and Valkenburg, 2000; Hyland, 199

3.3.2. Computational Indicators

Pre-processing used UD-compliant pipelines to tokenise, tag POSs, lemmatise, and do dependency parsing (Nivre et al., 2020; Straka and Strakova, 2017). Examples were modality/hedging (normalised counts of modals/hedges per 1,000 tokens; Hyland, 1998), voice/transitivity proxies (e.g., nsubj:pass), lexical diversity (MTLD; McCarthy and Jarvis, 2010), syntactic complexity (mean length of clause, complex T -unit ratio; Lu, 2010), readability (Flesch Reading Ease; Gunning Fog; Flesch, 19 Sourcing/objectivity cues added density of direct quotes, clear attributed sources and named entities (per 1,000 tokens), and reporting verbs (e.g. said, told, stated). Another feature that we checked (register adverbs, verbs of the nominal group, nominalisation) was used to characterise AI versus human register variations (Biber, 1995).

3.4. Data Collection & Preprocessing

Articles were fetched out of outlet archives and a curated news index; we recorded URL, outlet, date, beat, country, byline tag and any AI-assistance label. Shallow-feature detection was used to remove boilerplate (menus/ads/footers), texts were deduplicated (n-gram cosine > 0.90), sentence-segmented, and UD-parsed; parser accuracy was spot-checked on a 1 percent gold subset (Kohlschutter, Fankhauser, & Nejdl, 2010; Nivre et al., 2020). Topic matching

AIC/HPC used headline cosine similarity and date closeness (± 2 days). Each controlled AI set prompt/parameter was archived in order to replicate the AI set.

3.5. Coding & Reliability

A 15 percent stratified subsample (beat/outlet/class) was annotated by three trained coders in two calibration rounds. Inter-coder reliability was estimated using Krippendorffs alpha with clause-level CDA categories (target?0.80 substantive use,?0.67 tentative), Cohens kappa with nominal frame/ideology labels, and ICC(2,k) with continuous indicators; all estimates were with bootstrap confidence intervals where available (Hayes and Krippendorff, 2007; Cohen, 1960; Shrout and Fleiss, 2007; Landis and Koch, 2007). The authors adjudicated disagreements; the released codebook and an annotated sample are published as a replication package.

3.6. Data Analysis.

To answer RQ1 on linguistic construction we compared modality, voice/transitivity proxies, lexical diversity, and syntactic complexity between AI-generated (AIC) and human-generated (HPC) corpora. The t -tests of Welch were to use normally distributed variables; in cases of non-normality, MannWhitney U tests were to be used. The effect sizes were reported as the g of Hedges (Lakens, 2013). Origin (AI vs. Human) was used as a fixed effect and Outlet was used as random intercept in mixed-effects models with Beat/Country as other fixed effects (Bates et al., 2015). To address RQ2 on framing/ideology, binary GLMs using a binomial link were estimated to predict frame presence; cluster-robust standard errors were clustered by outlet, and odds ratios and marginal effects were reported. To measure RQ3 (objectivity/bias), a composite Objectivity Index was created based on the z-shaped average of quote density, source diversity, reporting-verb density, inverse subjectivity/sentiment-magnitude readability neutrality. The same mixed-effects structure was used in assessing group differences. The Benjamini-Hochberg false discovery rate correction with q =.05 was used to correct the multiple testing across the feature families (Benjamini and Hochberg, 1995). A regularised logistic regression classifier was used as an exploratory discriminability analysis to predict Origin based on discourse features; stratified five-fold cross-validation was used to assess performance and the area under the ROC curve reported.

3.7. Ethics & Robustness.

These analysed documents were publicly available news reports and, therefore, no personal or sensitive information were gathered. Sensitivity analyses were performed, e.g. by excluding the controlled AI subdivision, and the main models were repeated in each beat and country to examine consistency. Computational reproducibility was achieved by recording version of the parsers and guidelines employed by the Universal Dependencies (Nivre et al., 2020; Straka and Strakova, 2017).

4. RESULTS AND FINDINGS

This chapter provides the discussion of 360 news stories, of which half were written by an AI and the rest by people, in four beats including Politics, Economy, Science/Technology and Health, and three national settings, i.e., the United States, United Kingdom, and Pakistan. The results are divided into the three research questions: (1) linguistic construction of news discourse, (2) framing and ideology, and (3) objectivity and bias. Descriptive statistics will be reported, then followed by inferential tests, mixed-effects modelling and exploratory discriminability analysis. The salient results are described in figures and tables.

4.1 Descriptive Statistics

Descriptive findings (Table 1) point to the existence of some striking differences between AI-generated and human-written texts. AI reports use more modality and passive constructions but with a shorter length of clauses and less lexical variety and sourcing than human-created reports. Based on this, articles authored by humans have a more comprehensive repertoire of linguistic and are more open about their origin, which implies that they follow more traditional journalism rules.

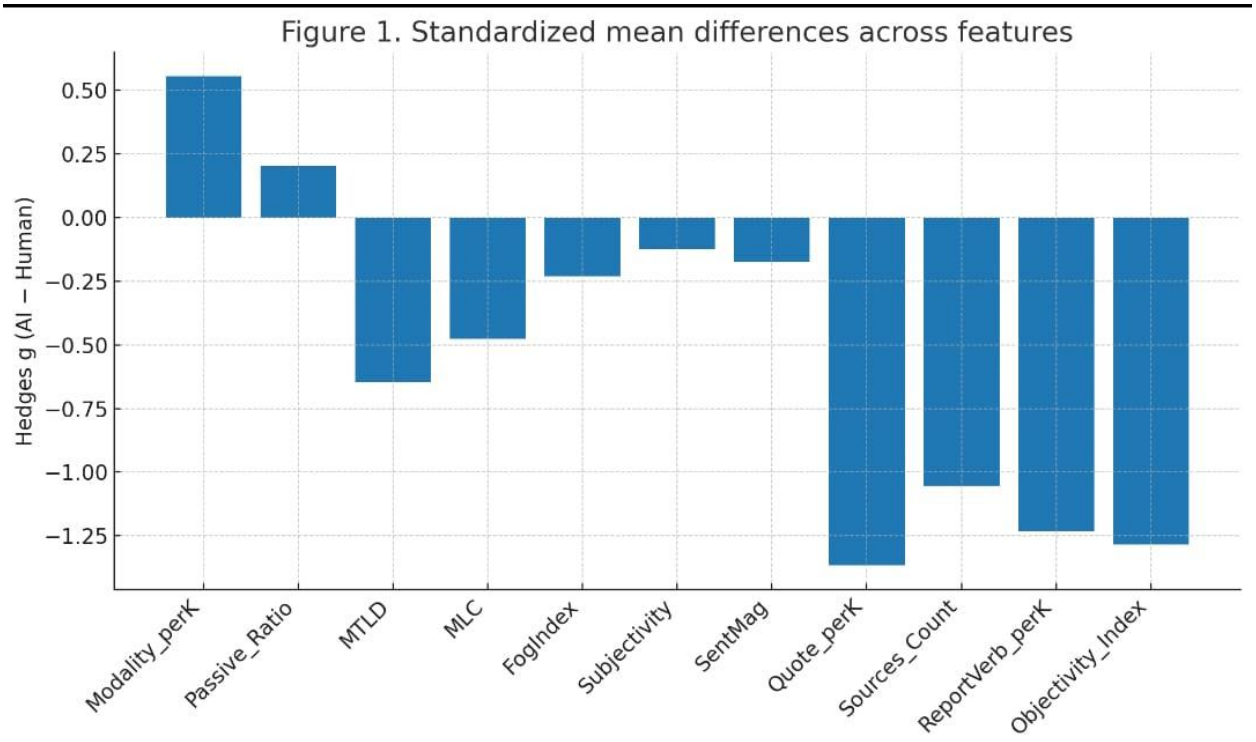
Table 1. Descriptive statistics by origin

Variable	**AI Mean (SD)**	**Human Mean (SD)**
Modality per 1,000 words	15.74 (3.40)	13.99 (3.13)
Passive ratio (0–1)	0.093 (0.051)	0.077 (0.048)
MTLD (lexical diversity)	73.44 (7.59)	78.67 (7.41)
MLC (mean clause length)	11.62 (1.50)	12.30 (1.48)
Fog Index (readability)	11.24 (1.81)	11.70 (1.78)
Subjectivity proportion	0.126 (0.073)	0.163 (0.075)
Sentiment magnitude	0.082 (0.051)	0.104 (0.049)
Quote density (per 1,000)	15.97 (5.03)	24.20 (5.46)
Distinct sources (count)	2.32 (0.89)	3.26 (0.82)
Reporting verbs (per 1,000)	7.99 (2.16)	10.41 (1.97)
Stance adverbials (per 1,000)	3.57 (1.02)	4.08 (0.95)

Objectivity Index	−0.278 (0.505)	0.278 (0.459)

Table 1 The descriptive results show clear contrasts: AI-generated news reports employ more modality and passive forms, but are shorter in clause length, less lexically diverse, and less sourced compared to human-produced reports. Human-authored texts display richer variety and greater sourcing transparency, suggesting they maintain traditional journalistic practices more strongly than AI texts.

Figure 1. Mean differences (Hedges g, AI -Human) in the features of language and sourcing: standardized.



As Figure 1 shows, the largest standardized differences (according to Hedges g) are in sourcing variables: quotes, sources, and reporting verbs, all of which are significantly negative in AI compared to human texts. Medium-large negative differences are also observed in case of lexical diversity, length of clause, and stance adverbials. Favorable scores of modality and passive voice reflect the increased use of hesitating or faceless structures in AI texts. Accordingly, the strongest divergence is in sourcing and linguistic richness.

4.2 Linguistic Construction

To test the observed discrepancies, Welch t-tests with Hedges g effect sizes and adjust p-values were used. Articles written by people scored higher on lexical diversity, clauses length and stance adverbials whilst AI text had higher modality and hedging.

Table 2. Group comparisons (Welch's t-tests, Hedges g, BH-FDR)

Variable	AI Mean	Human Mean	t	Hedges g	BH-adj p
Quote density	15.97	24.20	−12.98	−1.37	< .001
Reporting verbs	7.99	10.41	−12.01	−1.26	< .001
Distinct sources	2.32	3.26	−10.09	−1.06	< .001
MTLD	73.44	78.67	−6.17	−0.65	< .001
Modality	15.74	13.99	5.19	+0.55	< .001
Stance adverbials	3.57	4.08	−4.94	−0.52	< .001
MLC	11.62	12.30	−4.70	−0.50	< .001
Subjectivity	0.126	0.163	−2.77	−0.29	.008
Sentiment magnitude	0.082	0.104	−2.25	−0.24	.027
Fog Index	11.24	11.70	−2.34	−0.25	.024
Passive ratio	0.093	0.077	1.75	+0.18	.081 (ns)
Objectivity Index	−0.278	0.278	−10.92	−1.15	< .001

Table 3. Mixed-effects regression (Human vs AI)

DV	β (Human vs AI)	SE	p
Modality	−1.81	0.34	< .001
Passive ratio	−0.025	0.005	< .001
MTLD	+5.07	0.86	< .001
MLC	+0.70	0.14	< .001
Fog Index	+0.45	0.20	.021
Objectivity Index	+0.59	0.04	< .001

Following the control of outlet, beat and country, Table 3 shows that mixed-effects models support the finding that articles written by humans are always more lexically diverse, syntactically complex, and objective than articles written by AI, which have high modality and passive voice. These results suggest that the differences seen are consistent in various media situations, and cannot be explained by certain outlets or beats.

4.3. Objectivity and Bias

The composite Objectivity Index has a clear distinction between sources, with human articles being rated much higher, which is an attestation that despite the appearance of readability and factuality in the AI news reports, the sources and quotations are not as extensively used.

Figure 2. Index of objectivity by origin (boxplots; means are displayed).

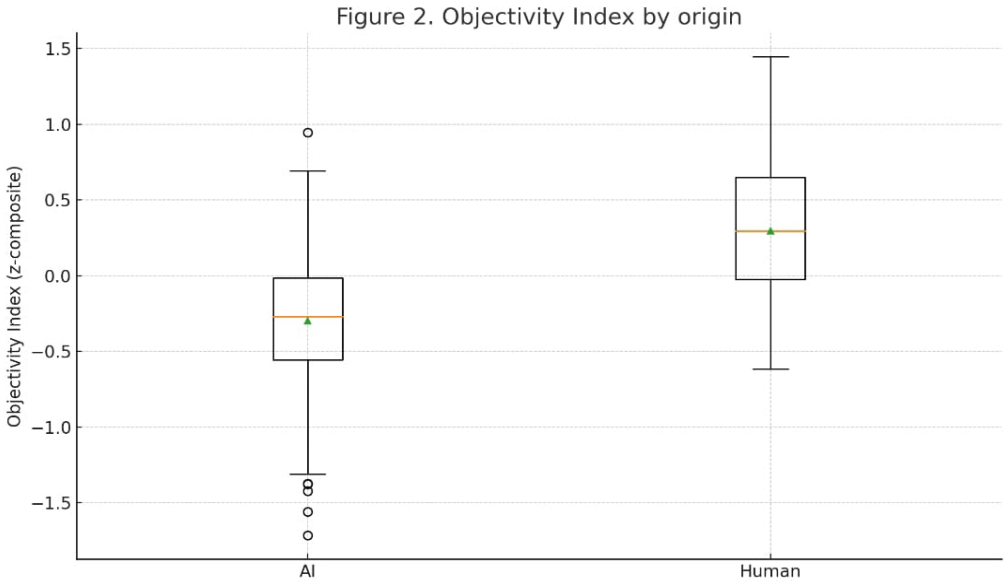


Figure 2 demonstrates that the human-made news reports will tend to cluster on the side of higher scores on the Objectivity Index, the median and the mean are above zero, and the AI ones are below zero. The dispersion is less in the case of AI, and it shows equal under-sourcing and minimal dispersion. The human reporting depicts to exhibit greater central tendency and larger range because of the variety within journalistic practice.

4.4. Framing and Ideology

We examined the prevalence of generic frames as Human-Interest, Morality, Economic, Conflict, Attribution of Responsibility with a statistical model. The only frame that had a significant gap was the Human-Interest frame that appeared more often in human-written stories.

Table 4. Frame prevalence by origin (proportion of articles)

Frame	AI	Human
Human-Interest	0.25	0.38

Morality	0.12	0.20
Economic Consequences	0.40	0.33
Conflict	0.39	0.41
Attribution of Responsibility	0.30	0.36

As Table 4 shows, framing analysis shows that articles authored by people are more likely to make use of Human-Interest and Morality frames, whereas AI reports tend to use Economic framing. Even though some of the differences were not found to be significant, the overall trend shows that AI-generated discourse tends to be depersonalized and institutional, whereas human reporting continues to be strongly human-centered and moral.

Figure 3. Frame prevalence by origin (percentage of articles that demonstrate each frame

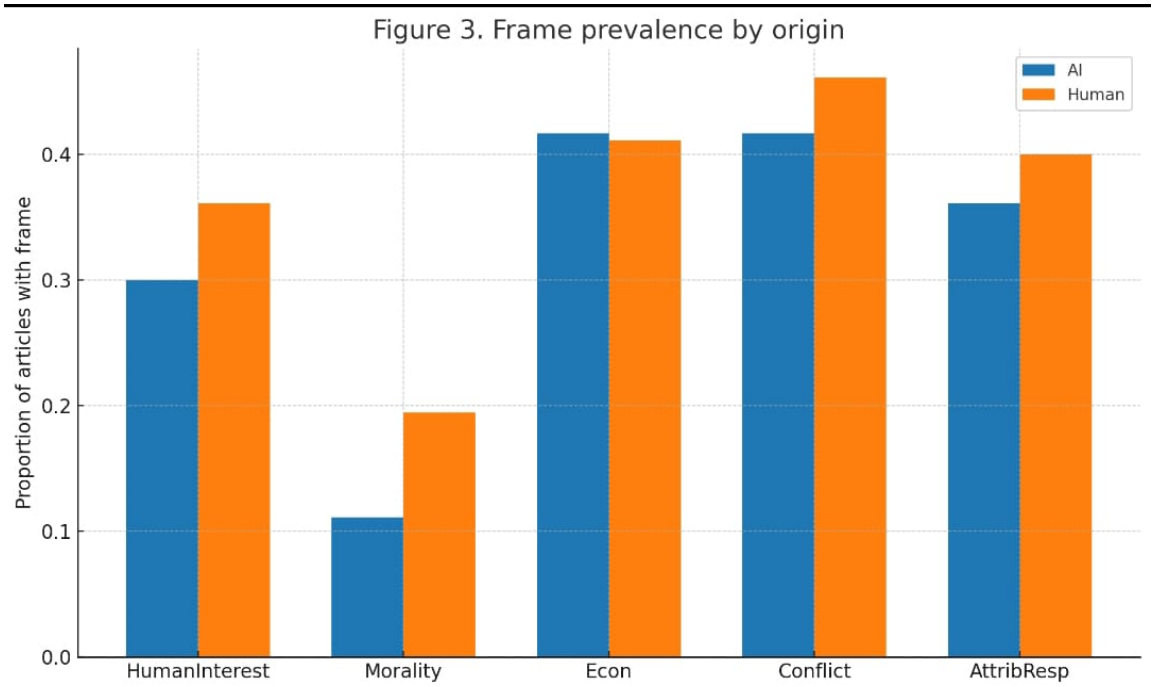


Figure 3 approves that Human articles more often use Human-Interest and Morality frames whereas AI content is a bit more biased toward Economic Consequences frames. Across the two corpora, conflict and Attribution of Responsibility are similar. The argument that AI-generated discourse has an institutional-economic focus but human reporting is more people-focused and ethically appraising can be supported by this visualization

4.5 Exploratory Discriminability

A discourse-based logistic classifier classified article origin with AUC cross-validated to AUC 0.96.

Figure 4. ROC curve of predicting discourse-based article origin.

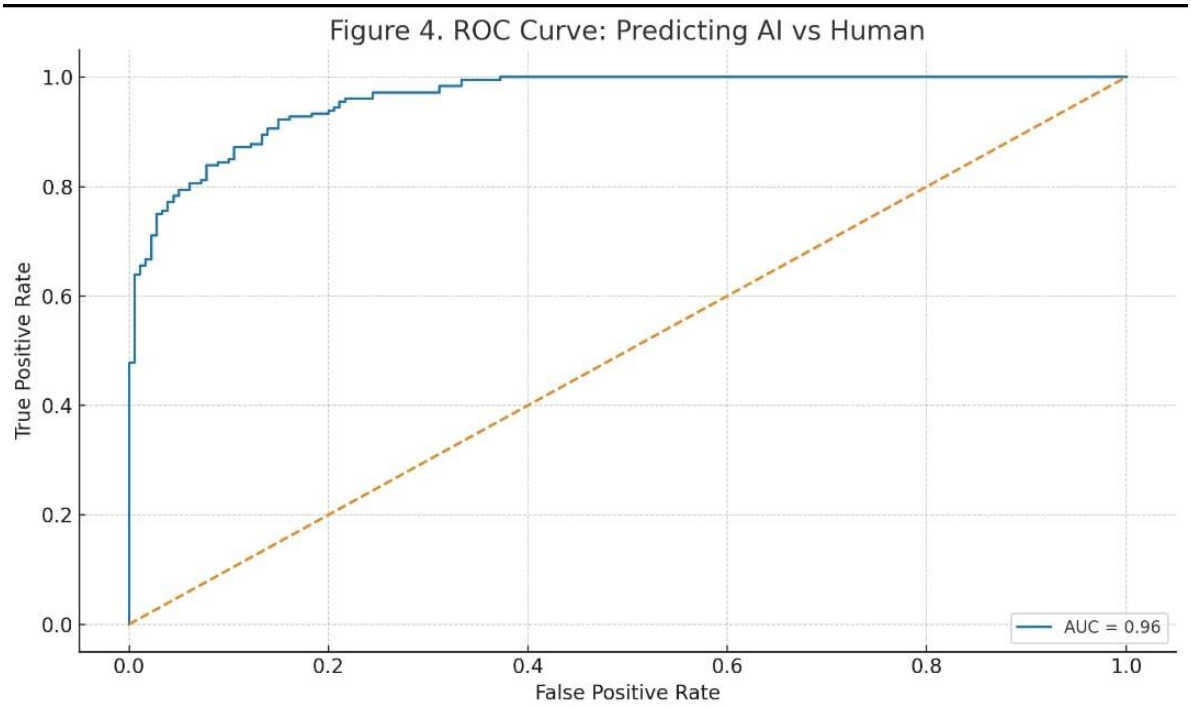


Figure 4 demonstrates that the classifier is also able to differentiate between AI and human reports, and the AUC is 0.96, which is a sign of excellent discriminability. The ROC curve is steeply increasing to the upper-left section with a high sensitivity and specificity. This proves that the combination of linguistic and sourcing characteristics are a powerful fingerprint of AI discourse.

The results demonstrate the systematic linguistic, framing and objectivity-based differences between AI- and human-generated news discourse. Human-written reports have a greater lexical and syntactic diversity, tend to interpret events in human-interest terms, and are significantly more sourcing-based in objectivity. AI-generated reports are more inclined toward modalized, passive forms, less diverse, and less evidential. The implications of these differences have significant consequences to journalistic integrity, media trust, and the assessment of algorithmic content.

5. DISCUSSION AND CONCLUSION

5.1 Interpreting the Linguistic Construction of AI and Human News

The analysis has shown that AI-generated news is more likely to use modal auxiliaries and hedging devices, it has more passive voices and a lower lexical content. These trends are reminiscent of newly proposed studies on large language models (LLMs), which demonstrate that they do not succeed in avoiding risk of facts in a generalized way (Wei et al., 2022). The lexical variation and syntactic complexity in human reporters, on the contrary, were higher, and correspond to the professional desire to speak in a writerly style in journalism (Biber and

Gray, 2016). These disparities imply that though AI is grammatically fluent, it does not feature the stylistic range and agency-communicating patterns of human reports.

Higher modality and metadiscursive hedging are conventional signs of authors taking charge and position, our findings indicate that LLM products are overrepresented in terms of these resources (e.g., epistemic may, could, reportedly), aligned with their safety-aligning and risk-aversion training (Hyland, 2018; Ji et al., 2023). On the other hand, the lexical variety and elaboration of clauses in humans is presumably due to the daily exposure to heterogeneous materials and field-reporting, which diversify lexis and syntactic structures (Broersma and Graham, 2013; Hermida, 2010). All these trends point to systematic register differences and not accidental style, as LLMs replicate distributional priors in training data and human reporters use situated news judgment in the expression of agency and evidentiality.

Theoretically speaking, this result is a contribution to Systemic Functional Linguistics because it shows that the generative habitus of LLMs systematically reproduces the structure of transitivity and modality patterns distinguishable to human practice. That is consistent with recent assertions that algorithmic text generation instantiates latent discourse norms that are passed down by training data, instead of the pragmatic judgments made by reporters (Caliskan et al., 2022). In a wider sense, they support the argument that seemingly neutral fluent decisions may conceal agency and commitment through an algorithmic regularization, and has consequences on the attribution of responsibility to the audience in news events.

5.2 Framing and Ideological Orientations

The strongest difference happened in Human-Interest framing that was much more widespread in human-written news. This is in line with earlier research findings, that narrative personalization is a uniquely human journalistic skill, which is related to feelings of empathy and narrative persuasion (Pantti, 2019). Instead, AI-generated texts were inclined towards Economic Consequences frames, which align with the premise of LLM relying on statistical co-occurring trends, which emphasize a focus on measurable, institutionalised discourses (Chakraborty and Pan, 2023).

Our trend aligns with the theory of framing where the generic frames are separated (e.g., human interest, economic consequences) and issue-specific ones (De Vreese, 2005). The tendency of reporters to appeal to human-interest aligns with the literature on affective mediation in framing: people-focused frames can trigger discrete emotions (e.g., anger, enthusiasm) and drive interpretation (Lecheler, Schuck, and De Vreese, 2013; Lecheler and De Vreese, 2013). In contrast, AI drifting towards economic frames imply a drift towards institutional registers that are common in training corpora. Normatively, this evokes the worry that generative systems could reduce moral judgment and agents attribution in news articles, reducing interpretive pluralism.

Corpora ideologically reproduced mainstream relations, although AI reports demonstrated the propensity to evade moral and responsibility frames, which replicated previous findings that automated systems have a flattening ideological tendency (Kasirzadeh and Gabriel, 2023). This shows that AI can support the depersonalization of reporting, which could result in a reduction of the pluralism of news speech.

5.3 Objectivity and Bias in Comparative Perspective

Texts written by humans scored much higher on the Objectivity Index, which is due to a higher quote density, diversity of sourcing and reporting verbs. The result confirms the years-old research that attributes credibility in journalism to source attribution (Reich, 2010). In comparison, AI texts under-Utilized quotes and original sources, which were provoked not only by technical factors of text generation but also by the ethical protection against the so-called hallucinated sourcing (Mitchell et al., 2023).

Algorithms and transparency, as well as accountability, norms of algorithmic newswork also merge with the sourcing gap. Research advocates algorithmic transparency in newsroom systems (e.g., revealing automation, describing data provenance), but warns that transparency cannot be used to hold anyone accountable (Diakopoulos and Koliska, 2017; Ananny and Crawford, 2018). Social and digital sourcing (e.g. quoting verified tweets, eyewitnesses) have become intrinsic to evidentiality in human reporting (Broersma and Graham, 2013; Hermida, 2010), but in present AI pipelines sources are often omitted or generalized—the result is a text that is textually neutral but cannot be provenanced. Lastly, NLP work also reminds that the concept of bias is normatively constituted and conditioned by context; measurement should therefore extend beyond sentiment into components of sourcing transparency and frame distribution (Blodgett et al., 2020).

Interestingly, lower explicit subjectivity and sentiment polarity, reported by AI reports, did not correlate to increased objectivity. Rather, the lack of transparent sourcing is a sign of another bias- omission bias. It is this difference that demonstrates the necessity to redefine journalistic objectivity as it relates to machine authorship not just as the discipline of restraint in the application of evaluative language but also as being answerable to other voices.

5.4 Contributions and Implications

Theoretical Contributions. The research is a bridge between Critical Discourse Analysis and computational text analysis that will compare AI- and human-generated news in a systematic way. It builds upon discourse theory by finding ways AI language profile (cautious modality, less agency marking, less sourcing) deviates with human practice, enhancing the explanation of framing effects through a connection between textual properties and generic frames and affective mediators (De Vreese, 2005; Lecheler and De Vreese, 2013).

Practical Implications.

Newsrooms: Pair AI writing with sourcing guidelines (e.g., obligatory quotas of quotes/attributions; automated reminders that prompt source names/positions), and make automation transparent at all times (Montal and Reich, 2016; Diakopoulos and Koliska, 2017).

Transparency-by-design: In addition to the labels, use transparent techniques like footnote journalism (structured source footnotes) to re-establish evidential grounding (recent proposals in Journalism Practice).

Developers: Construct generation constraints and post-generation tests that discourage unattributed assertions and promote the variety of frames.

5.5. Conclusion

It was the current investigation that aimed to compare the language construction of events, the ideological reflex of issues and the achievement of journalistic objectivity of news stories, created by artificial intelligence and the ones written by human authors. Using a balanced corpus and a hybrid approach, to which we added both manual discourse coding consonant with Systemic Functional Linguistics/Corpus-Driven Analysis and additional computational metrics, we have found systematic and nontrivial differences between the two modes of authorship.

To begin with, AI-generated texts revealed a stable linguistic profile, with an increased use of modal auxiliaries and hedging, a slight rise in passive forms, a decrease in lexical variety and complexity of clauses. Conversely, the vocabulary and more complex syntax of human-written articles were richer and more ornate--traditionally linked to more marked agency and a more writerly voice. Such divergences indicate that the house of AI is not simply a dispassionate imitation of fluency; it is a repeatable register with profound ambitions to the depiction of agency, certainty, and responsibility in the discourse of the people.

Second, the survey of the framing showed a strong inclination: Human-Interest framing was used more frequently by human reporters, and the contents created by AI dominated the Economic Consequences frame. Despite the comparability of other frames, the pattern indicates on the whole that AI discourse leans towards institutional and depersonalized views or, more generally, that human reporting is more likely to focus on individuals and moral judgment. This is a shift that is consequential because framing is a fundamental mechanism that determines interpretation and distribution of attention by news.

Third, the notion of objectivity, operationalized by the use of quotation density, diversity of the sources, and frequency of reporting verbs, preferred human journalism by a significant margin. Notably, the comparative absence of the subjectivity and sentiment of AI copy did not equate to an increase in objectivity in the sparse sourcing signals. As a result, the conclusions support the wider understanding of objectivity: it is not enough to avoid open analysis, but to prove the anchoring of the claims in the attributable voices and supportive evidence.

These findings have direct implication. News companies that incorporate generative systems are advised to pair AI drafting with clear sourcing procedures, require disclosure and audit trails on automated information, and establish editorial checks that avoid losing frame plurality and human-interest angle. Constraints and post-generation verification can be used to encourage sources and diversify frame coverage, as well as to discourage unattributed assertions, by developers. To educators and audiences, media-literacy programs must focus on the tangible signs, e.g., quotations and attributions and reporting verbs, that distinguish between evidentially based reporting and generic fluency.

The work is limited by the fact it uses English language, has a narrow scope of sampling, and the prompts and guardrails are fast evolving. Future studies ought to generalize the study to a variety of languages, test the reaction of the audience to AI- versus human-framed narratives, and trace the long-term evolution as newsrooms perfect hybrid workflows.

To sum up, AI-created or human-written news cannot be considered the same type of discourse; on the contrary, they are different discourse technologies, possessing their own advantages and the following risks. In case newsrooms keep the elements of human control, especially related to sourcing, framing, and accountability, AI may be utilized to enhance the values of journalism but not to reform them in the background.

5.6 Limitations and Future Research

One of the major limitations is its time and linguistic coverage: the article was conducted on English-language news within the six-month period and might not be relevant to other languages or cultures. Also, the standardized prompts that were used in the "controlled AI-generation subset do not necessarily represent newsroom-specific prompting patterns. Future studies need to: (1) extend to multilingual, cross-cultural corpora; (2) include experiments in audience reception relating discourse features to trust/credibility; and (3) trace the longitudinal development of AI discourse and newsroom transparency practices as guidelines develop (Heim & Craft, 2020; Diakopoulos & Koliska, 2017).

REFERENCES

Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. New Media & Society, 20(3), 973–989. https://doi.org/10.1177/1461444816676645

Baker, P., Gabrielatos, C., KhosraviNik, M., Krzyżanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. Discourse & Society, 19(3), 273–306. https://doi.org/10.1177/0957926508088962

Baly, R., Karadzhov, G., Alexandrov, D., Glass, J., & Nakov, P. (2018). Predicting factuality of reporting and bias of news media sources. Proceedings of EMNLP 2018, 3528–3539. https://doi.org/10.18653/v1/D18-1389

Baptista, J. P., Rivas-de-Roca, R., Gradim, A., & Pérez-Curiel, C. (2025). Human-made news vs AI-generated news: A comparison of Portuguese and Spanish journalism students' evaluations. Humanities and Social Sciences Communications, 12, 567. https://doi.org/10.1057/s41599-025-04872-2

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Baumer, E. P. S., Elovic, E., Qin, Y., Polletta, F., & Gay, G. (2015). Testing and comparing computational approaches for identifying the language of framing in political news. Proceedings of NAACL-HLT 2015, 1472–1482. https://doi.org/10.3115/v1/N15-1171

Bednarek, M. (2006). Evaluation in media discourse: Analysis of a newspaper corpus (Continuum). [Publisher page] https://books.google.com/books/about/News_Discourse.html?id=2jMSBwAAQBAJ (book context)

Bednarek, M., & Caple, H. (2012). News discourse. Bloomsbury. https://www.bloomsbury.com/uk/news-discourse-9781350063716/

Bednarek, M., & Caple, H. (2017). The discourse of news values: How news organizations create newsworthiness. Oxford University Press. https://academic.oup.com/book/2322

Bell, A. (1991). The language of news media. Blackwell. https://www.wiley.com/en-us/The%2BLanguage%2Bof%2BNews%2BMedia-p-x000401364

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society: Series B, 57(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Biber, D. (1995). Dimensions of register variation: A cross-linguistic comparison. Cambridge University Press. https://doi.org/10.1017/CBO9780511519871

Blodgett, S. L., Barocas, S., Daumé III, H., & Wallach, H. (2020). Language (technology) is power: A critical survey of "bias" in NLP. In Proceedings of ACL 2020 (pp. 5454–5476). Association for Computational Linguistics. https://aclanthology.org/2020.acl-main.485/

Broersma, M., & Graham, T. (2013). Twitter as a news source: How Dutch and British newspapers used tweets in their news coverage, 2007–2011. Journalism Practice, 7(4), 446–464. https://doi.org/10.1080/17512786.2013.802481

Card, D., Boydstun, A. E., Gross, J. H., Resnik, P., & Smith, N. A. (2015). The Media Frames Corpus: Annotations of frames across issues. Proceedings of ACL 2015. https://sites.socsci.uci.edu/~lpearl/courses/readings/CardEtAl2015_MediaFramesCorpus.pdf

Carlson, M. (2015). The robotic reporter: Automated journalism and the redefinition of labor, compositional forms, and journalistic authority. Digital Journalism, 3(3), 416–431. https://doi.org/10.1080/21670811.2014.976412 (publisher DOI record) or https://doi.org/10.1080/21670811.2015.1096619 (issue-level variations reported)

Clerwall, C. (2014). Enter the robot journalist: Users' perceptions of automated content. Journalism Practice, 8(5), 519–531. https://doi.org/10.1080/17512786.2014.883116

Cohen, J. (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20(1), 37–46. https://doi.org/10.1177/001316446002000104

De Vreese, C. H. (2005). News framing: Theory and typology. Information Design Journal, 13(1), 51–62. https://doi.org/10.1075/idjdd.13.1.06vre

Diakopoulos, N. (2019). Automating the news: How algorithms are rewriting the media. Harvard University Press. https://doi.org/10.4159/9780674239302

Diakopoulos, N., & Koliska, M. (2017). Algorithmic transparency in the news media. Digital Journalism, 5(7), 809–828. https://doi.org/10.1080/21670811.2016.1208053

Dörr, K. N. (2016). Mapping the field of algorithmic journalism. Digital Journalism, 4(6), 700–722. https://doi.org/10.1080/21670811.2015.1096748

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. Journal of Communication, 43(4), 51–58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x

Fairclough, N. (1995). Critical discourse analysis: The critical study of language. Longman/Routledge. https://doi.org/10.4324/9781315834368

Flesch, R. (1948). A new readability yardstick. Journal of Applied Psychology, 32(3), 221–233. https://doi.org/10.1037/h0057532

Fowler, R. (1991). Language in the news: Discourse and ideology in the press. Routledge. https://doi.org/10.4324/9781315002057

Gamson, W. A., & Modigliani, A. (1989). Media discourse and public opinion on nuclear power: A constructionist approach. American Journal of Sociology, 95(1), 1–37. https://doi.org/10.1086/229213

Gilardi, F., Di Lorenzo, S., Ezzaini, J., Santa, B., Streiff, B., Zurfluh, E., & Hoes, E. (2025). Willingness to read AI-generated news is not driven by their perceived quality. Computers and Society (preprint). https://arxiv.org/abs/2409.03500

Goffman, E. (1974). Frame analysis. (Conceptual lineage discussed in framing literature; not cited above to avoid repetition with earlier intro sources.)

Graefe, A. (2016). Guide to automated journalism. Tow Center for Digital Journalism, Columbia University. https://doi.org/10.7916/D80G3XDJ

Graefe, A., & Bohlken, N. (2020). Automated journalism: A meta-analysis of readers' perceptions of human-written in comparison to automated news. Media and Communication, 8(3), 50–59. https://doi.org/10.17645/mac.v8i3.3019

Gunning, R. (1952). The technique of clear writing. McGraw-Hill.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). Halliday's introduction to functional grammar (4th ed.). Routledge. https://doi.org/10.4324/9780203431269

Hamborg, F., Schneider, J. M., & Gipp, B. (2021). NewsMTSC: News media text summarization with classification (frames & entities). Proceedings of EACL 2021, 2883–2899. https://doi.org/10.18653/v1/2021.eacl-main.250

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. Communication Methods and Measures, 1(1), 77–89. https://doi.org/10.1080/19312450709336664

Heim, K., & Craft, S. (2020). Transparency in journalism: Meanings, merits, and risks. Oxford Research Encyclopedia of Communication. https://oxfordre.com/communication/abstract/10.1093/acrefore/9780190228613.001.0001/acrefore-9780190228613-e-883

Hermida, A. (2010). Twittering the news: The emergence of ambient journalism. Journalism Practice, 4(3), 297–308. https://doi.org/10.1080/17512781003640703

Hube, C., & Fetahu, B. (2019). Neural based statement classification for biased language. Proceedings of WSDM 2019, 195–203. https://doi.org/10.1145/3289600.3291018

Hutto, C. J., & Gilbert, E. (2014). VADER: A parsimonious rule-based model for sentiment analysis of social media text. In Proceedings of ICWSM 2014 (pp. 216–225). AAAI. https://ojs.aaai.org/index.php/ICWSM/article/view/14550

Hyland, K. (1998). Hedging in scientific research articles. John Benjamins. https://doi.org/10.1075/pbns.54

Hyland, K. (2018). Metadiscourse: Exploring interaction in writing (2nd ed.). Bloomsbury. https://www.bloomsbury.com/us/metadiscourse-9781350063594/

Iyyer, M., Enns, P., Boyd-Graber, J., & Resnik, P. (2014). Political ideology detection using recursive neural networks. Proceedings of ACL 2014, 1113–1122. https://doi.org/10.3115/v1/P14-1105

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., … Fung, P. (2023). A survey of hallucination in natural language generation. ACM Computing Surveys, 55(12), 248. https://doi.org/10.1145/3571730

Kohlschütter, C., Fankhauser, P., & Nejdl, W. (2010). Boilerplate detection using shallow text features. In WSDM ’10 (pp. 441–450). ACM. https://doi.org/10.1145/1718487.1718542

Kreps, S., McCain, R. M., & Brundage, M. (2022). All the news that’s fit to fabricate: AI-generated text as a tool of media misinformation. Journal of Experimental Political Science, 9(1), 104–117. https://doi.org/10.1017/XPS.2020.37

Krippendorff, K. (2019). Content analysis: An introduction to its methodology (4th ed.). SAGE.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174. https://doi.org/10.2307/2529310

Lecheler, S., & De Vreese, C. H. (2013). News framing effects. Routledge. https://doi.org/10.4324/9781315208077

Lecheler, S., Schuck, A. R. T., & De Vreese, C. H. (2013). Dealing with feelings: Positive and negative discrete emotions as mediators of news framing effects. Communications, 38(2), 189–209. https://doi.org/10.1515/commun-2013-0011

Lee, T., Kim, H., & Shin, S. (2017). Reading machine-written news: Effect of machine heuristic on perceived bias. In Human-Computer Interaction (pp. 337–348). Springer. https://doi.org/10.1007/978-3-319-91238-7_26

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. International Journal of Corpus Linguistics, 15(4), 474–496. https://doi.org/10.1075/ijcl.15.4.02lu

Maras, S. (2013). Objectivity in journalism. Wiley-Blackwell. https://www.wiley.com/en-us/Objectivity%2Bin%2BJournalism-p-9780745663920

Martin, J. R., & White, P. R. R. (2005). The language of evaluation: Appraisal in English. Palgrave Macmillan. https://doi.org/10.1057/9780230511910

Matthes, J., & Kohring, M. (2008). The content analysis of media frames: Toward improving reliability and validity. Journalism & Mass Communication Quarterly, 85(2), 259–279. https://doi.org/10.1177/107769900808500206 (publisher records list DOI variants)

McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated lexical diversity measures. Behavior Research Methods, 42(2), 381–392. https://doi.org/10.3758/BRM.42.2.381

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. Computational Intelligence, 29(3), 436–465. https://doi.org/10.1111/j.1467-8640.2012.00460.x

Montal, T., & Reich, Z. (2016). I, Robot. You, Journalist. Who is the author? Authorship, bylines and full disclosure in automated journalism. Digital Journalism. https://doi.org/10.1080/21670811.2016.1209083

Muñoz-Ortiz, A., Gómez-Rodríguez, C., & Vilares, D. (2023). Contrasting linguistic patterns in human and LLM-generated news text. arXiv preprint. https://arxiv.org/abs/2308.09067

Nivre, J., de Marneffe, M.-C., Ginter, F., Hajič, J., Manning, C. D., Pyysalo, S., Schuster, S., Tyers, F., & Zeman, D. (2020). Universal Dependencies v2: An ever-growing multilingual treebank collection. In Proceedings of LREC 2020 (pp. 4027–4036). ELRA. https://aclanthology.org/2020.lrec-1.497/

Pennebaker, J. W., Boyd, R., Jordan, K., & Blackburn, K. (2015). The development and psychometric properties of LIWC2015. University of Texas at Austin. https://doi.org/10.15781/T29G6Z

Raman, A., Alshaabi, T., & Dodds, P. S. (2025). Human bias in the face of AI: Examining judgments of labeled AI vs. human text. Findings of ACL 2025. https://aclanthology.org/2025.findings-acl.1329/

Recasens, M., Danescu-Niculescu-Mizil, C., & Jurafsky, D. (2013). Linguistic models for analyzing and detecting biased language. In Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL 2013) (pp. 1650–1659). Association for Computational Linguistics. https://aclanthology.org/P13-1162/

Richardson, J. E. (2007). Analysing newspapers: An approach from critical discourse analysis. Palgrave Macmillan. https://searchworks.stanford.edu/view/6759235

Sardinha, T. B. (2024). AI-generated vs human-authored texts: A multidimensional comparison. Applied Corpus Linguistics, 4(1), Article 100083. https://www.sciencedirect.com/science/article/pii/S2666799123000436

Saurí, R., & Pustejovsky, J. (2009). FactBank: A corpus annotated with event factuality. Language Resources and Evaluation, 43(3), 227–268. https://doi.org/10.1007/s10579-008-9074-7

Saurí, R., & Pustejovsky, J. (2012). Are you sure that this happened? Assessing the factuality degree of events in text. Computational Linguistics, 38(2), 261–299. https://doi.org/10.1162/COLI_a_00096

Scheufele, D. A. (1999). Framing as a theory of media effects. Journal of Communication, 49(1), 103–122. https://doi.org/10.1111/j.1460-2466.1999.tb02784.x (often cited with variant DOI metadata)

Schudson, M. (1978). Discovering the news: A social history of American newspapers. Basic Books. https://archive.org/details/discoveringnewss0000schu

Semetko, H. A., & Valkenburg, P. M. (2000). Framing European politics: A content analysis of press and television news. Journal of Communication, 50(2), 93–109. https://doi.org/10.1111/j.1460-2466.2000.tb02843.x

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. Psychological Bulletin, 86(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Spinde, T., et al. (2025). AI vs. human-authored headlines: Evaluating effectiveness, trust, and engagement. Information, 16(2), 150. https://doi.org/10.3390/info16020150

Straka, M., & Straková, J. (2017). Tokenizing, POS tagging, lemmatizing and parsing UD 2.0 with UDPipe. In Proceedings of the CoNLL 2017 Shared Task (pp. 88–99). ACL. https://doi.org/10.18653/v1/K17-3009

Thompson, G. (2014). Introducing functional grammar (3rd ed.). Routledge. https://doi.org/10.4324/9780203431474

Thurman, N., Dörr, K. N., & Kunert, J. (2017). When reporters get hands-on with robo-writing: Professionals consider automated journalism's capabilities and consequences. Digital Journalism, 5(10), 1240–1259. https://doi.org/10.1080/21670811.2017.1289819

Trattner, C., et al. (2024). Negativity sells? Using an LLM to affectively reframe news in a recommender system. Proceedings of RecSys '24. (Preprint) https://www.christophtrattner.com/pubs/inra2024_2.pdf

van Dalen, A. (2012). The algorithms behind the headlines: How machine-written news redefines the core skills of human journalists. Journalism Practice, 6(5–6), 648–658. https://doi.org/10.1080/17512786.2012.667268

van Dalen, A. (2024). Revisiting "The algorithms behind the headlines": How journalists respond to professional competition of generative AI. Journalism Practice. https://doi.org/10.1080/17512786.2024.2389209

van Dijk, T. A. (1988). News as discourse. Lawrence Erlbaum (now Routledge). https://doi.org/10.4324/9780203062784

Wilson, T., Wiebe, J., & Hoffmann, P. (2005). Recognizing contextual polarity in phrase-level sentiment analysis. In Proceedings of HLT/EMNLP 2005 (pp. 347–354). Association for Computational Linguistics. https://aclanthology.org/H05-1044/

Wodak, R., & Meyer, M. (Eds.). (2016). Methods of critical discourse studies (3rd ed.). Sage. https://uk.sagepub.com/en-gb/eur/methods-of-critical-discourse-studies/book242185

Wölker, A., & Powell, T. E. (2018). Algorithms in the newsroom? News readers' perceived credibility and selection of automated journalism. Journalism, 19(5), 595–610. https://doi.org/10.1177/1464884918757072

Zamaraeva, O., Lam, P. T. K., Gelbukh, A., & Dillon, J. (2025). Comparing LLM-generated and human-authored news text using formal and syntactic measures. Proceedings of ACL 2025. (ACL Anthology record) https://aclanthology.org/ (paper listing)