The *Zé Lensky* Dataset: A Brazilian Portuguese Twitter Corpus for Russo-Ukraine War Stance and Sentiment Analysis

Andreis G. M. Purim[1], **Kārlis Kuškēvics**[2]

[1]Institute of Computing - University of Campinas (UNICAMP)

[2]Ukrainian Congress Committee of America (UCCA)

andreispurim@proton.me, kkuskevics@ucca.org

Abstract. *This paper presents a work-in-progress corpus of over 200,000 Brazilian Portuguese tweets related to the Russia-Ukraine war, collected between 2022 and 2025. The dataset includes metadata and annotations for stance and irony, with particular attention to partisan and culturally specific expressions such as "zé lensky", which we hypothesize indicate processes of sociolinguistic enregisterment in partisan speech. This resource enables further exploration of political discourse in Brazilian social media and supports future studies on narrative dynamics, user behavior, and informal political language. The dataset is available on GitHub*[1].

1. Introduction

The Russian full-scale invasion of Ukraine, which began in 2022, generated extensive global activity on social media. X (formerly Twitter), in particular, served not only as a platform for time-sensitive discourse but also as a window into public opinion and informal commentary, offering perspectives beyond institutional news reporting. Since then, many *Natural Language Processing* (NLP) studies have focused on English-language content, and while other languages have been considered, English remains the predominant medium of analysis, with X/Twitter serving as a primary channel for public discourse. Several NLP tasks have been applied to examine social media reactions to the conflict. Sentiment analysis is used to quantify emotional tone, while stance detection identifies alignment with specific entities or propositions (e.g., support for Ukraine or Russia). Other approaches include misinformation classification, named entity recognition, and event extraction. These methods have largely been benchmarked on English corpora, with limited adaptation to other linguistic settings. Among the less frequently addressed languages, Brazilian Portuguese represents a particularly relevant case. Brazil has one of the largest X/Twitter-using populations in the Global South, yet its digital discourse about the conflict has not been systematically studied in the NLP literature.

A particular interest of this research is the use of irony, memes, informal rhetorical style, and partisan derogatory terms observed in segments of the Brazilian public discourse - which relates to the sociolinguistic phenomena of *indexicality* and *enregisterment*. Derogatory and colloquial expressions such as *"zé lensky"* (to refer to President

[1]https://github.com/AndreisPurim/ZLDB. In compliance with X's Content Redistribution Policy, we release only tweet IDs and derived metadata; rehydration and annotation details are in the README. We use public data (no individual consent sought), release only tweet IDs/metadata, avoid PII in text or supplements, and take care not to reinforce stereotypes or amplify harmful narratives. Researchers may contact the first author for access or guidance.

Zelensky) and *"putinetes"* (for pro-Putin sympathizers) illustrate the presence of culturally specific and implicitly charged language. These phenomena present important challenges for automated stance and sentiment analysis in Brazilian Portuguese. This work-in-progress addresses these gaps by constructing a new, dedicated corpus of over 200,000 Brazilian Portuguese tweets related to the Russia-Ukraine war from 2022 to 2025. Moreover, a subset of 24,000 tweets was annotated, each being labeled with a stance (ranging from strongly pro-Russia to strongly pro-Ukraine) and a level of irony. This annotated subset enables both quantitative and qualitative exploration of online war discourse in Brazil.

2. Related Work

Russia-Ukraine war datasets: several large-scale datasets have been created for NLP research on the Russia-Ukraine war, focusing primarily on stance detection, sentiment analysis, and misinformation tracking. Many of these datasets concentrate on English or use multilingual translation pipelines. For example, a comprehensive corpus by Lamprou et al. comprises over 127 million tweets in 70 languages [Lamprou et al. 2024], while the "Tweets in Time of Conflict" dataset offers 500 million tweet IDs [Chen and Ferrara 2022]. Other relevant tweet-based datasets include the "Sunset Conflict" corpus with approximately 70 million tweets[2], and the annotated "UKRUWAR22" dataset with 55,000 multilingual tweets[3]. The "CrisisHateMM" dataset integrate posts from X/Twitter, Facebook, and Reddit, labeled for hate speech severity and intent [Bhandari et al. 2023]. Finally, the "Unveiling Global Narratives" corpus provides 1.5 million multilingual tweets from official media accounts, annotated for stance and sentiment [Hakimov and Cheema 2024], though translations into English may obscure culturally specific expressions.

Portuguese-language stance and irony detection: research on stance detection and irony classification in Brazilian Portuguese has advanced in recent years. A benchmark dataset for zero-shot stance detection in Portuguese was introduced and evaluated using strategies adapted from English, showing strong performance even against supervised models [Pavan and Paraboni 2024]. The "UstanceBR" corpus provides a large multimodal dataset of Brazilian X/Twitter posts annotated for stance, supporting baseline experiments with transformer-based models [Pereira et al. 2023]. In parallel, the IDPT shared task offered benchmark corpora of tweets and news headlines for systematic evaluation of irony detection methods [Corrêa et al. 2021]. A dedicated irony dataset for Portuguese social media was also compiled to support supervised learning approaches and subsequent benchmarks [Marten and Freitas 2021]. Together, these works have established important resources and modelling strategies for stance and irony detection tailored to Brazilian Portuguese.

Sociolinguistic indexicality and enregisterment: the use of specific lexical items as markers of political stance or group identity is often analyzed with the concepts of *indexicality* and *enregisterment*. Indexicality describes how particular terms point to social meanings and ideological alignments beyond their literal content [Silverstein 2003]. Enregisterment refers to the process by which such terms (nicknames, slogans, or

[2]Sunset Conflict corpus: `https://www.kaggle.com/dsv/8003074`
[3]UKRUWAR22 corpus: `https://dx.doi.org/10.21227/g0me-wa71`

slurs) become socially recognized as part of a group-specific vocabulary [Gal 2019, Hoinkes 2019]. In political discourse, adopting or rejecting an enregistered term signals alignment or opposition, making these lexical items interesting features for stance and irony detection. Irony in this context often functions as an off-record critique, and digital platforms accelerate the circulation of such enregistered terms, reinforcing their role as *shibboleths* in online communities [Zappavigna 2022, Squires 2010]. In Brazilian politics, clear examples of enregisterment include the "papo reto" register [Silva 2022] and the slurs "mortadela" (for supporters of the Workers' Party) and "coxinha" (for opponents of the Workers' Party) [Bittencourt 2016]. In this dataset, the use of *"zé lensky"* may serve as an enregistered term for indexing political alignment.

3. Methodology and Results

As a work-in-progress, we continue to expand coverage, refine filters, and incrementally label stance and irony. As of writing, the corpus contains over 200,000 Portuguese-language posts (Jan 2022–Jan 2025), collected with a `lang:pt` constraint. We combined general conflict queries over canonical entities (e.g., "Ucrânia", "Rússia", "Putin", "Zelensky") with a targeted partisan/colloquial lexicon that captures culturally specific forms often used ironically (e.g., "zé lensky", "putinetes"). To reduce amplification bias and redundancy, we excluded pure retweets and de-duplicated near-identical texts[4]. Heuristics and metadata filters were applied to exclude tweets originating from outside Brazil, based on user location strings, time zones, and geotag hints. Nevertheless, the dataset might still include some tweets from Portugal and other Portuguese-speaking regions.

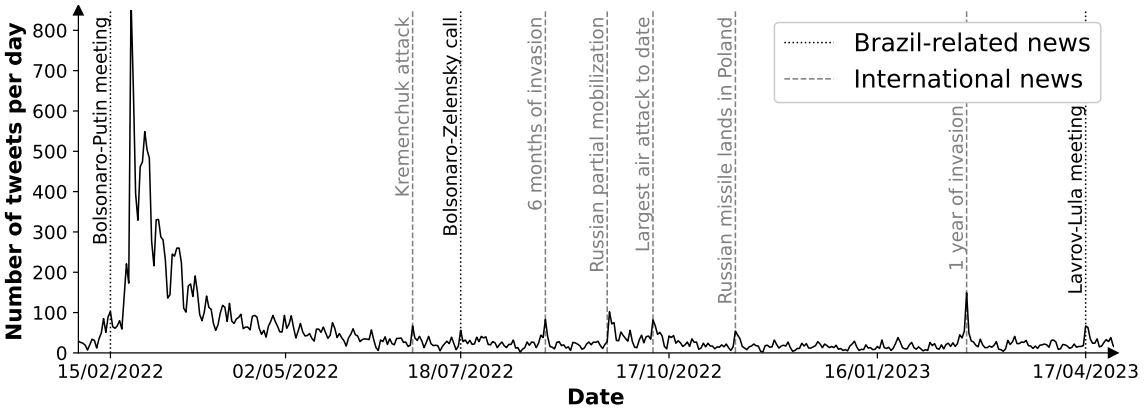


Figure 1. Tweet count of Brazilian news sources per day with key events.

As an initial step, we marked tweets from verified and official Brazilian news outlets (e.g., *Folha de S.Paulo*). This allowed us to analyze official media cycles across the dataset's timeframe. Figure 1 presents the daily tweet count for this subset, annotated with key geopolitical events. The data indicates peaks aligned with major international developments, such as Russia's partial mobilization announcement or the one-year anniversary of the invasion. Brazil-specific diplomatic events, such as Bolsonaro's visit to Moscow shortly before the war or the 2023 meeting between Russian Foreign Minister Lavrov and President Lula, are reflected in the dataset but show comparatively lower volume.

To explore how public discourse evolved over time outside media accounts, we selected user posts from two comparable time windows: Week 13 of 2022 and Week 13 of

[4]The repository README provides the search queries, time bounds, and preprocessing scripts.

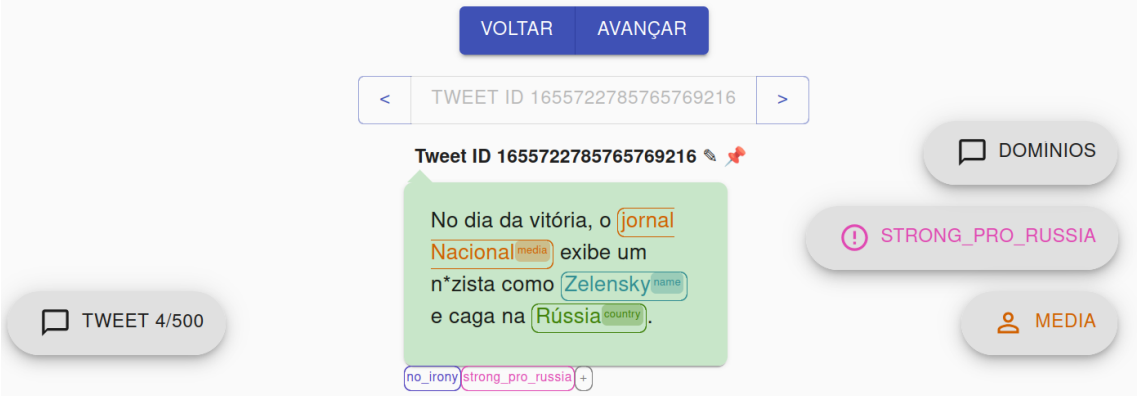


Figure 2. Example of one tweet annotated in the Assis tool.

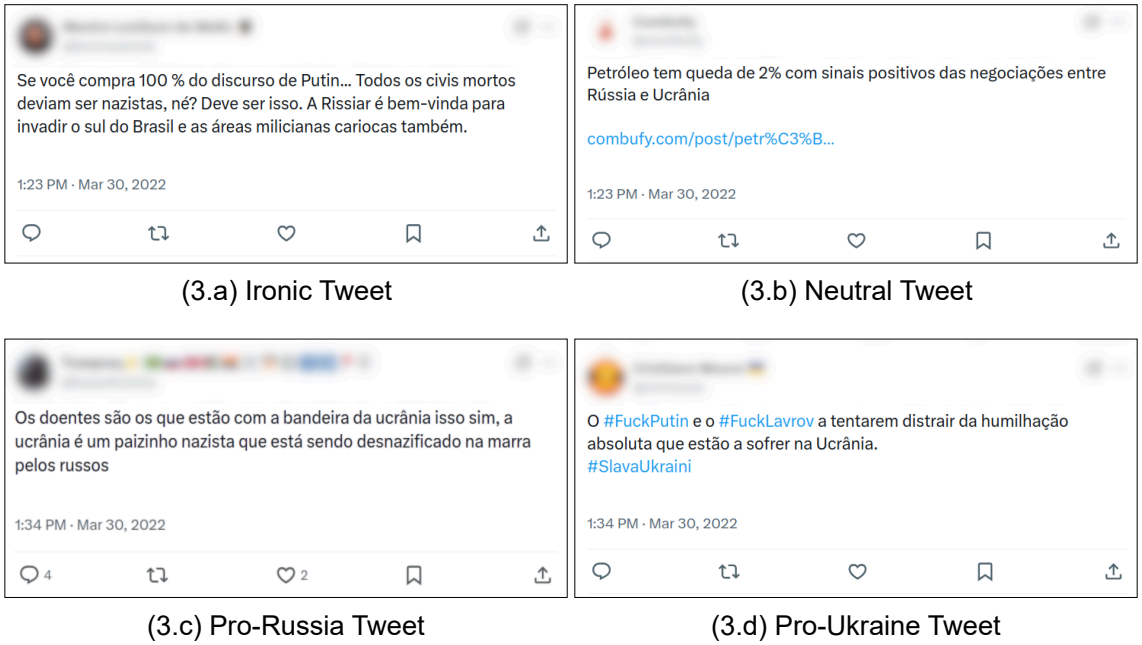
2024 (approximately the last week of March in both years). These two samples were arbitrarily drawn from the broader dataset to enable a controlled temporal comparison - they fall near the anniversary of the invasion (around week 9) and contain a stable volume of discourse (approximately 12,000 posts each), enabling comparison while minimizing the impact of overlapping domestic political events. The data was annotated using a modified version of the *Assis Annotation Tool*[5] [Foerste et al. 2023, Purim and dos Reis 2023].

First, the data was processed by a GPT-4o annotation backend which pre-classified each tweet along two dimensions: (i) **stance**, on a scale from -2 (strongly pro-Russia) to $+2$ (strongly pro-Ukraine); and (ii) **irony**, ranging from 0 (no irony) to 2 (high irony). As seen in Figure 2, the model outputs were manually reviewed by a fluent Portuguese speaker familiar with regional slang and political discourse. Labels were manually revised to correct misinterpretations of sarcasm, idiomatic expressions, or political references. As future work, we will report a small inter-annotator agreement study (Cohen's κ / weighted κ) on a random sample to quantify labeling reliability.

The labeling scheme defines strong expressions of support or criticism (e.g., explicit endorsement of military actions, praise or condemnation of political leaders) as ± 2, while milder alignment (e.g., indirect sympathies or concerns) was labeled as ± 1. Tweets expressing opinions without favoring a side were labeled as neutral (0). Irony was assessed based on linguistic cues such as tone, sarcasm, figurative phrasing, or exaggeration. While this labeling process offers a functional approximation of discourse patterns, it remains inherently limited. Irony, in particular, is context-dependent and often difficult to classify reliably in short texts. Similarly, stance may be implied rather than stated, introducing ambiguity in edge cases. Future work will aim to improve annotation consistency, expand the benchmark set, add a multi-annotator pass, and report inter-annotator agreement (e.g., Cohen's κ, weighted κ / Krippendorff's α) on a random sample. Figure 3 presents four example tweets from the labeled dataset, illustrating: (3.a) high irony (irony $= 2$); (3.b) neutral stance (stance $= 0$); (3.c) strongly pro-Russia (stance $= -2$); and (3.d) strongly pro-Ukraine (stance $= +2$).

Figure 4 compares the distribution of stance across the two weekly samples in item (4.a). We observe a higher share of pro-Russia tweets in 2024 and a lower share of explicitly pro-Ukraine tweets relative to 2022. Because this comparison spans only two week-of-year snapshots, we treat it as descriptive rather than a trend; an extensive tem-

[5] Available at: `https://github.com/AndreisPurim/AssisAnnotationTool`

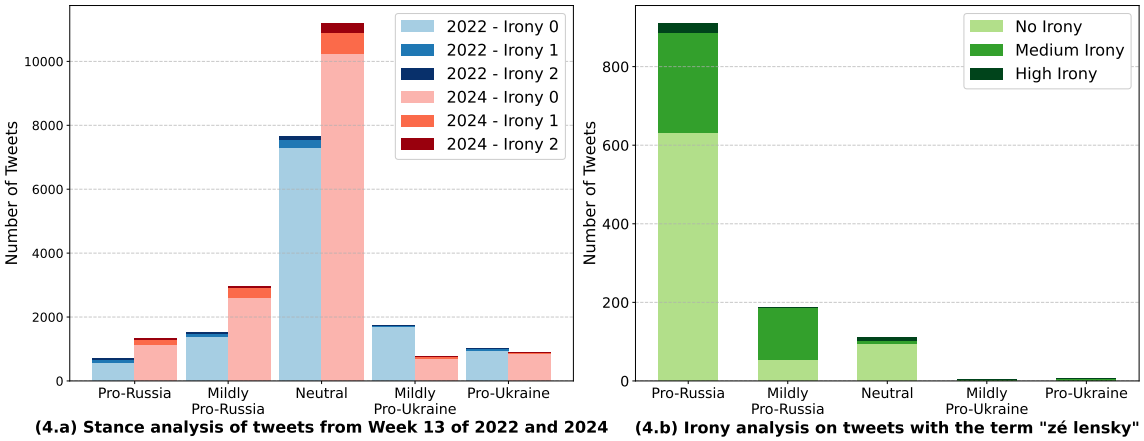


(3.a) Ironic Tweet (3.b) Neutral Tweet

(3.c) Pro-Russia Tweet (3.d) Pro-Ukraine Tweet

Figure 3. Examples of four classified tweets

poral analysis is necessary to determine whether this is a trend. Moreover, the relatively high presence of irony in tweets labeled as neutral may reflect the difficulty of accurately classifying ironic content, which often obscures clear political alignment and challenges both automated and human annotation. Additionally, as seen in (4.b), certain recurring lexical patterns, such as the ironic nickname *"zé lensky"* were more frequently observed in tweets classified as both pro-Russian, with a noticeable increase in irony (compared to the all tweets without 'zé lensky' in the same week).



(4.a) Stance analysis of tweets from Week 13 of 2022 and 2024 **(4.b) Irony analysis on tweets with the term "zé lensky"**

Figure 4. Preliminary stance and and Irony analysis.

The analyses presented above are preliminary and reflect an initial exploration of the dataset. While they highlight patterns in stance, irony, and rhetorical expression, a more comprehensive understanding will require deeper temporal, linguistic, and behavioral analysis. As the dataset expands and the annotation process matures, we will release updated scripts and a versioned dataset card documenting known limitations, sampling, and intended use.

3.1. Conclusion and Next Steps

This study is part of an ongoing effort to document and analyze Brazilian Portuguese discourse on X/Twitter about the Russia-Ukraine war. By labeling tweets for stance and irony, and separating media accounts from general users, we aim to provide a resource that helps capture how political attitudes and rhetoric evolve in a specific cultural context. Our initial comparison between 2022 and 2024 suggests differences not only in alignment but also in tone and expression; given the narrow temporal slice, we treat these as descriptive rather than trend claims. Future work will address:

- **Temporal coverage.** We plan to extend the manually annotated set to additional weeks (both event-aligned and ordinary) to test whether the 2022 vs. 2024 difference in stance and irony persists rather than reflect a single moment.
- **Annotation reliability.** We plan to add a small multi-annotator comparison (200–300 tweets) and report inter-annotator agreement (Cohen's κ for stance; weighted κ/Krippendorff's α for irony), with a brief error analysis of ironic cases. In addition, we may revisit and formalize the stance/irony labeling protocol (clarifying thresholds, adding borderline examples, and specifying adjudication rules) and, if warranted, adjust the scales to improve reliability.
- **Brazil filter precision.** One gap is the inaccuracy of the Brazil-focused filter. We plan to audit its precision on a random sample and report 95% confidence intervals, and to design a more robust filter complemented by a PT-BR vs. PT-PT lexical probe; procedures and estimates will be documented in the repository.
- **Portuguese AI model comparison.** We plan to benchmark the GPT-4o pipeline against Portuguese pretrained models (e.g., BERTimbau) on the annotated subset, evaluating both zero/few-shot and fine-tuned variants. We will report accuracy and F1 for stance and macro-F1 for irony, with confusion matrices for error analysis. Moreover, the Portuguese pretrained models can be fine-tuned with the annotated data and evaluated. This will also include a simple active-learning loop: a model scores the unlabeled pool and flags the most uncertain (or diverse) tweets for human review (e.g., via uncertainty sampling); the new labels are then added to the training set and the model is retrained, iterating to improve accuracy with fewer total annotations.
- **User profiles, enregisterment, topics, and amplification.** We plan to add post and user metadata (when available) to characterize user profiles and analyze group-level stance dynamics over time; apply topic modeling and clustering to identify and track salient themes; and examine engagement patterns (e.g., retweets, likes, replies) and basic diffusion measures. Together, these steps will help document how partisan nicknames and cues function as indexical markers of group identity (enregisterment) within Brazil's Russia–Ukraine discourse.

4. Funding, Acknowledgments, and Disclosure

This research received no external funding; all costs (e.g., API access, computers, and storage) were covered by the authors. We thank Ambassador Alda Vanaga (Embassy of Latvia in Brazil) and the Baltic Security Foundation (BSF), in particular Otto Tabuns and Olevs Nikers, for their constructive support during the preparation of this work. The views expressed are those of the authors and do not necessarily reflect the positions of UNICAMP, UCCA, or any other organization. The authors report no conflicts of interest.

References

Bhandari, A., Shah, S. B., Thapa, S., Naseem, U., and Nasim, M. (2023). Crisishatemm: Multimodal analysis of directed and undirected hate speech in text-embedded images from russia-ukraine conflict. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1994–2003.

Bittencourt, R. N. (2016). A culinária da política: coxinha, caviar e mortadela. *Revista Espaço Acadêmico*, 16(182):45–55.

Chen, E. and Ferrara, E. (2022). Tweets in time of conflict: A public dataset tracking the twitter discourse on the war between ukraine and russia. arXiv:v2, 2023-04-10.

Corrêa, U. B., Coelho, L., Santos, L., and de Freitas, L. A. (2021). Overview of the idpt task on irony detection in portuguese at iberlef 2021. *Procesamiento del Lenguaje Natural*, page 269–276.

Foerste, H. T. S., Purim, A. G. M., Souza, R. R., and Dos Reis, J. C. (2023). Assis: Online semi-automatic dialog annotation tool. In *Proceedings of the XIX Brazilian Symposium on Information Systems*, SBSI '23, page 37–44, New York, NY, USA. Association for Computing Machinery.

Gal, S. (2019). Making registers in politics: Circulation and ideologies of linguistic authority. *Journal of Sociolinguistics*, 23(5):450–466.

Hakimov, S. and Cheema, G. S. (2024). Unveiling global narratives: A multilingual twitter dataset of news media on the russo-ukrainian conflict. In *Proceedings of the 2024 International Conference on Multimedia Retrieval*, ICMR '24, page 1160–1164, New York, NY, USA. Association for Computing Machinery.

Hoinkes, U. (2019). Indexicality and enregisterment as theoretical approaches to the sociolinguistic analysis of romance languages. Technical report, Kiel University (MACAU Repository). Accessed 2025-04-07.

Lamprou, I., Shevtsov, A., Antonakaki, D., Pratikakis, P., and Ioannidis, S. (2024). Exploring crisis-driven social media patterns: A twitter dataset of usage during the russo-ukrainian war. In *Social Networks Analysis and Mining: 16th International Conference, ASONAM 2024, Rende, Italy, September 2–5, 2024, Proceedings, Part I*, page 70–85, Berlin, Heidelberg. Springer-Verlag.

Marten, G. S. and Freitas, L. A. d. (2021). The construction of a corpus for detecting irony and sarcasm in portuguese / a construção de um corpus para detectar a ironia e o sarcasmo em português. *Brazilian Journal of Development*, 7(5):47973–47984.

Pavan, M. C. and Paraboni, I. (2024). A benchmark for portuguese zero-shot stance detection. *Journal of the Brazilian Computer Society*, 30(1):469–479.

Pereira, C., Pavan, M., Yoon, S., Silva, R. M., Costa, P., Cavalheiro, L., and Paraboni, I. (2023). Ustancebr: A social media language resource for stance prediction.

Purim, A. G. M. and dos Reis, J. C. (2023). Active Learning for Natural Language Data Annotation. Technical Report IC-PFG-23-55, Institute of Computing, University of Campinas.

Silva, D. N. (2022). Papo reto: The politics of enregisterment amid the crossfire in rio de janeiro. *Signs and Society*, 10(2):239–264.

Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language & Communication*, 23(3):193–229. Words and Beyond: Linguistic and Semiotic Studies of Sociocultural Order.

Squires, L. (2010). Enregistering internet language. *Language in Society*, 39(4):457–492.

Zappavigna, M. (2022). Social media quotation practices and ambient affiliation: Weaponising ironic quotation for humorous ridicule in political discourse. *Journal of Pragmatics*, 191:98–112.