

Reported speech in social media: disentangling source and commitment

Marie Boscaro & Valentin Tinarrage

Institut Jean Nicod CNRS/ENS/EHESS/PSL

Abstract

Question and scope. This work addresses the question of the notional category of reported speech by looking at social networks and Twitter in particular. We ask what its manifestation is and how it differs from direct speech. We cast this question in the framework of the study of evidentiality, building on Aikhenvald’s view ([1],[2],[3]). We show that the boundaries between those two are blurred on social media and newly provide empirical arguments to this claim. We will finally explore the theoretical consequences of this blurring by considering assertions vs. reported speech within a general framework of veridicality (see [4]) and more broadly speech act theory ([8], [12]).

Methodology. Our study uses an annotated corpus of 11000 tweets related to emergency situations ([6]). We have annotated 1790 tweets for evidentiality and achieved a substantial agreement with a kappa coefficient of approximately 0.63. While in the NLP literature, evidentiality has mostly focused on eyewitnesses ([11], [9], [5]) and testimonies ([13], [14]), we have built our classification on Aikhenvald ([1],[2],[3]) and Willett ([10]). Our categories are described in Table 1 (see Table 2 for the numerical distribution). The main tripartition points to a distinction of enhanced degrees of commitment via reliability. Indeed, while “rumor” is a type of reported evidence, its nature as a poorly reliable source leads us to classify it together with lack of testimony. As shown in Table 2, the “relayed” category is dominant.

Speech Type	Category	Description
Direct	Direct	Testimony with first-hand evidence (photographs and videos).
Reported	Indirect	Second-hand evidence from a trusted source within a small social circle.
	Relayed	Information reported from a source outside of the speaker’s circle, mainly press or institutional sources (hyperlinks to news articles, tweets, etc.).
Unclear	Don’t Know	Testimonies with informational value, but the source is difficult to determine.
	Supposition	Information deduced from logical reasoning, general knowledge.
	Rumor	Vague information regarding the crisis, the source is unspecified or impersonal.
	No testimony	When the message doesn’t relay a testimony.

Table 1: Evidential categories for social media

Reported vs. Direct speech. Reported speech on Twitter takes on multiple forms, including more traditional ones, like a quoted citation with its source mentioned like in (1). Specific to online spaces, directly sharing hyperlinks to a news article, to another tweet, etc. provides the addressees with both the informational content as well as its source, as in (2). Moreover, when citing tweets, the “retweet” and “quote tweet” features remove the need for adding a link altogether, as the context is now automatically provided.

We observe that in the case of bare retweets (where a user shares another tweet without adding anything), it can be implied that the author endorses whatever information or opinion they’re sharing. The addition of a commentary allows the user to express their personal feelings and opinions on the situation, to add additional information, to voice their lack of endorsement, etc. as in (3). Lack of introduction of the quotational content, levels the sharp distinction between reported speech and direct assertions. Indeed, while retweeting (reported speech), the speaker might imply that they endorse the content entirely. The overwhelming use of reported speech in social media points to the fact that the speaker acts as if they endorsed the content, while signalling that they are not the main source, thus breaking the usual connection between strength of the source and strength of the commitment.

We will spell out the analysis of this phenomenon in a gradable theory of veridicality [4]. While it is normally assumed that assertions convey full commitment (in virtue of direct evidence) and reported speech lack of commitment (in virtue of indirectness), we will elaborate a notion of *non-committal assertivity* for cases where the assertion relies on second-hand sources.

Appendix

Table 2: Annotation Results

Category	Count	%
Relayed	1013	56.59%
Don’t Know	280	15.64%
Direct	217	12.12%
No testimony	167	9.33%
Supposition	113	6.31%
Total	1790	100.00%

- (1) *le maire de Rouen dit que la fumee n’est pas toxique alors que l’usine Lubrizol est classé SEVESO* <https://t.co/AREwqVeEPf>
(Translation: the mayor of Rouen says that the smoke is not toxic even though the Lubrizol plant is classified SEVESO <https://t.co/AREwqVeEPf>)
- (2) *Un incendie impressionnant s’est déclaré cette nuit à Rouen, à l’usine #Lubrizol qui fabrique des additifs pour lubrifiants, le site est classé ”Seveso”: récit de @YannBouchery à 6h dans #RTLPetitMatin* <https://t.co/1mfAnSvV65> <https://t.co/OnpfViO5We>
(Translation: An impressive fire broke out last night at the Lubrizol factory in Rouen, which manufactures lubricant additives. The site is classified as ”Seveso”. @YannBouchery reports on the fire at 6am on #RTLPetitMatin <https://t.co/1mfAnSvV65> <https://t.co/OnpfViO5We>)
- (3) *gistupa DIRECT - Incendie à Rouen : le feu à l’usine Lubrizol dangereux ? Images, infos, par @martin_lmr : <https://t.co/11qmApRpiz> via @LInternauteInfo. Tt com tchernobyl on viendra ns dire que les vents tournent au dessus du nuage! Respirer cette saloperie est dangereux au long terme.*
(Translation: gistupa DIRECT - Fire in Rouen: Is the fire at the Lubrizol factory dangerous? Images and information by @martin_lmr: <https://t.co/11qmApRpiz> via @LInternauteInfo. Just like Chernobyl we will be told that the winds are blowing above the cloud! Breathing this crap is dangerous in the long run.)

References

- [1] Aikhenvald, A. Y. (2004). Evidentiality. Oxford University Press.
- [2] Aikhenvald, A. Y., & Dixon, R. M. W. (2014). The grammar of knowledge: A cross-linguistic typology. Oxford University Press.

- [3] Aikhenvald, A. Y. (2018). *The Oxford handbook of evidentiality*. Oxford University Press.
- [4] Giannakidou, A., & Mari, A. (2020). *Truth and veridicality in grammar and thought: Mood, modality and propositional attitudes*. The University of Chicago Press.
- [5] Haider, S., Azhar, M., Khatoon, S., Alshamari, M., & Afzal, M. (2022). Automatic classification of eyewitness messages for disaster events using linguistic rules and ML/AI approaches. *Applied Sciences*.
- [6] Kozłowski, D., Lannelongue, E., Saudemont, F., Benamara, F., Mari, A., Moriceau, V., & Boumadane, A. (2020). A three-level classification of French tweets in ecological crises. *Information Processing & Management*.
- [7] Portner, P. (2009). *Modality*. Oxford University Press.
- [8] Portner, P. (2018). *Mood*. Oxford University Press.
- [9] Truelove, M., Vasardani, M., & Winter, S. (2015). Towards credibility of micro-blogs: Characterising witness accounts. *Geojournal*.
- [10] Willett, T. (1988). A cross-linguistic survey of the grammaticization of evidentiality. *Studies in language*.
- [11] Zahra, K., Imran, M., & Ostermann, O. F. (2020). Automatic identification of eyewitness messages on Twitter during disasters. *Information Processing and Management*.
- [12] Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge University Press.
- [13] Sorin Adam Matei Olteanu, Carlos Castillo, Fernando Diaz, and Emma S. Spiro. What to Expect When the Unexpected Happens: Social Media Communications Across Crises. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, pages 994–1009, 2015.
- [14] Imran, M., Elbassuoni, S. M., Castillo, C., Diaz, F., & Meier, P. (2013). Extracting information nuggets from disaster-related messages in social media. In *Proceedings of the 10th International ISCRAM Conference* (pp. 1-10). May.