



HAL
open science

Evidence type and trustworthiness

Marie Boscaro, Anastasia Giannakidou, Alda Mari

► To cite this version:

Marie Boscaro, Anastasia Giannakidou, Alda Mari. Evidence type and trustworthiness. Susana Rodríguez Rosique; Jordi M. Antolí Martínez. *The Grammar of Interaction. Epistemicity, information management and discourse in language use*, 46 (4), John Benjamins Publishing Company, pp.98-131, 2025, IVITRA Research in Linguistics and Literature, 9789027244352. <10.1075/ivitra.46.04bos>. <halshs-05475973>

HAL Id: halshs-05475973

<https://shs.hal.science/halshs-05475973v1>

Submitted on 25 Jan 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

Evidence type and trustworthiness

The view from social media

Marie Boscaro | Institut Jean Nicod CNRS

Anastasia Giannakidou | University of Chicago

Alda Mari | Institut Jean Nicod CNRS

 <https://doi.org/10.1075/ivitra.46.04bos>

 Available under a CC BY-NC-ND 4.0 license.

Pages 98–131 of

The Grammar of Interaction: Epistemicity, information management and discourse in language use

Edited by Susana Rodríguez Rosique and Jordi M. Antolí Martínez

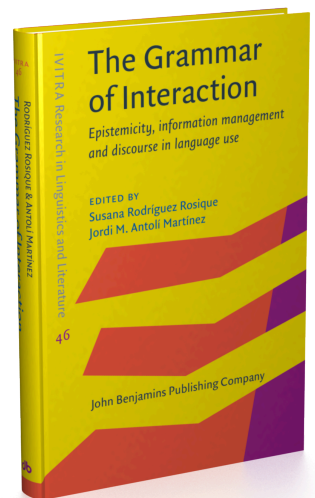
[*IVITRA Research in Linguistics and Literature*, 46]

2025. xix, 348 pp.

© John Benjamins Publishing Company

This electronic file may not be altered in any way. For any reuse of this material, beyond the permissions granted by the Open Access license, written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

For further information, please contact rights@benjamins.nl or consult our website at benjamins.com/rights



Evidence type and trustworthiness

The view from social media

Marie Boscaro¹, Anastasia Giannakidou² & Alda Mari¹

¹ Institut Jean Nicod CNRS | ² University of Chicago

This paper provides a first analysis of the relationship between evidentiality and the speaker's veridical commitment (Giannakidou and Mari 2021c) in social networks. Our corpus study highlights that speakers ground their assertive statements in various types of evidence, including indirect evidence (Willett 1988). Our statistical results indeed indicate a strong correlation between Reported evidence and bare Assertions, suggesting that speakers often rely on reported information while still expressing a strong veridical commitment. We demonstrated that this should not be interpreted as a non-cooperative linguistic behavior, even if indirect evidence is typically considered inadequate for a strong veridical commitment. We argued that this perspective is too restrictive and that evidence type should be decoupled from reliability. We propose that the perceived adequacy of an evidence type stems from a subjective evaluation of trustworthiness of the speaker of both the evidence and the propositional content it conveys. This evaluation is shaped by the speaker's informational basis, which includes subjective preferences, prior beliefs, background assumptions and biases, as well as external knowledge about the world.

Keywords: evidence, trustworthiness, assertion, social networks, speech acts

1. Assertion, evidence and veridical commitment

The linguistic category of evidentiality has a venerable tradition in linguistic theorizing. Careful cross-linguistic work has brought to light robust grammatical systems where languages encode source of evidence – and while the phenomenon is not homogeneous cross-linguistically (with some languages lacking altogether such a grammatical system), it is accurate to say that the speech act of assertion relies on the veridical commitment of the speaker to the content of the utterance (Searle 1975; Harnish 1994; Portner 2018; Giannakidou and Mari 2021a, a.o.),

and this in turn depends on the type of evidence the speaker has. It has also been acknowledged that only *best possible evidence* constitutes a solid ground for veridical commitment in bare assertion. Typically, when the speaker has witnessed an action herself, she takes that action to be true and reports it in an unqualified assertion, such as *Bill plays the piano (I saw him play)*; but if the speaker has not witnessed the act of playing the piano she may report this as *Bill plays the piano, I hear*. In this case the speaker relies on indirect, or reported evidence, and many languages actually *have to* encode the difference in every sentence. In the assertion that relies on reported evidence, the speaker may have reservations about the truth of the statement, and the assertion may at first glance appear ‘weaker’. These distinctions between evidence type and their associated reliability, and their interactions with veridical commitment, have been fixed in scales of preferred evidence (see among others De Haan 1999; Faller 2002; McCready 2015).

In this paper, we address the relation between the type of evidence – direct, indirect, or no evidence at all – , and the degree of veridical commitment for the assertion as observed in a new discourse area, X (formerly known as Twitter).¹ We tackle this question from the point of view of veridical commitment, evidentiality research, and philosophical literature on testimony epistemology. Our main research questions are the following: what types of evidence do Twitter users tend to rely on more frequently? What are the interactions between evidence-type and veridical commitment in this digital context? The empirical study extends on a corpus of over 3,137 tweets annotated for speech acts and evidence type categories. Our analysis addresses the question of the statistical correlations between speech act categories and evidence-type categories and discusses the theoretical consequences of the correlations found in our empirical survey.

We achieve two main results. First, we provide a linguistic-based classification of evidence in online discourses. Second, we show that the category of indirect evidence correlates strongly with assertion on Twitter – defying the expectation we outlined at the beginning of current linguistic theories that treat indirect evidence as ‘weaker’ (i.e. conveying weaker veridical commitment). The strong correlation between reported evidence and assertions that we observed in social media shows that this evidence type can be associated with a high degree of trustworthiness, and enhances an assertive speech act on Twitter. We propose a theoretical account of this correlation, arguing that in social media contexts, reported evidence can still function as best possible grounds for full veridical commitment. This perspective calls for a reconsideration of traditional assumptions about evi-

1. As this work was carried on a corpus that was gathered before the advent of X, we will continue to refer to the social network with its previous appellation Twitter.

dence reliability and more specifically about indirect evidence which is typically analyzed as a bad candidate for justifying strong veridical commitment. In this regard, we introduce the concept of *trustworthiness evaluation*, proposing that evidence type and proximity do not solely determine the perceived trustworthiness of evidence, or its assumed intrinsic strength (as it has been previously assumed by Matthewson 2010); instead, it is shaped by a subjective evaluation on the part of the speaker, based on his or her informational state.

Before we move on, we want to provide some background on the relation between assertion, veridical commitment and evidence, which will be necessary if we want to understand what issues are at stake. When speakers make assertions or assess statements of others they form veridicality judgments about the uttered content, and these judgments depend on what speakers know or believe to be true based on the evidence they have, on what kinds of expectations arise given the evidence or other more subjective preferences (Giannakidou and Mari 2021c). Paul Grice in his classic paper *Logic and Conversation* established Quality (*Do not say that for which you lack evidence*) as one of the foundational principles of cooperative conversation. Quality is formulated as a negative principle in Grice, a filter on what *not* to say: do not say that which you know to be false, do not say that for which you lack adequate evidence. Building on Grice's seminal work and the long-standing tradition that this work has initiated (a.o. Bach and Harnish 1984; Gunlogson 2001; Farkas and Bruce 2010; Lauer 2013; Krifka 2015), Giannakidou and Mari (2021c,a) formulate Quality positively as the Veridicality Principle of Assertion below, which posits *veridical commitment* as the basis for assertion:

(1) Principle of Veridicality of Assertion

A sentence *S* can be asserted by a speaker *A* if and only if *A* is veridically committed to the content π of *S*, i.e., if and only if *A* knows or believes π to be true.

The Veridicality principle can be understood as a sincerity condition of cooperative conversation, and veridical evidence relies on knowledge and (rational) belief. Knowing or believing rationally that a content is true relies on having evidence for that content, and on judging the evidence as reliable. This notion of reliable evidence will require a measure of what we call *trustworthiness* and introduce later in this paper.

Knowledge and evidence-based belief establish full veridical commitment. When the speaker is fully committed to the truth of the content conveyed, by asserting *It is raining*, she signals that she has some evidence that she evaluates as reliable that it is raining, and intends the proposition *p* 'It is raining', to be added to the common ground of the conversation (Stalnaker 1978), thus becoming common knowledge. This function of adding a proposition to the common ground is taken to be the core function of assertion in every formal analysis of asser-

tion. (see a.o. Beyssade and Marandin 2009; Farkas and Bruce 2010; Ginzburg 2012; Krifka 2015; Geurts 2019; Yalcin 2024). When the speaker lacks reliable evidence she will use some device, say an epistemic modal, to indicate uncertainty: for instance, she might utter *It might be raining* if she is guessing; or she may use a *biased* modal (Kratzer 1991; Pietrandrea 2005; Portner and Partee 2008; Giannakidou and Mari 2018) *It must be raining* if she has some evidence but is still not absolutely sure, e.g., if she sees people coming in with wet umbrellas. This MUST sentence still carries uncertainty, but there is preference (yet not knowledge) towards the prejacent proposition *It is raining*. In this case, what is added to the common ground is the modal sentence, not *p* itself. If a speaker relies on inadequate evidence but does not use a modal, her statement can be analyzed as a case of an uncooperative assertion or deception.

Several theoreticians (see e.g., De Haan 1999; Faller 2002; Murray 2010; Ricci and Rossari 2018, to cite just a few) generally accept that the more reliable the source of evidence the stronger the evidence and the commitment are. To the best of our knowledge, all theories of evidentiality distinguish between direct and indirect (reported) evidence. If reported evidence is used, a quantity implicature can be triggered that the commitment of the speaker is lower, perhaps because they didn't witness the event and can't verify the source. This is an implicature indeed, because it is cancelled when the reported evidence is attributed to an epistemic authority: e.g. *The weather app says it will rain tomorrow* will trigger confidence in the prediction because the weather app is an epistemic authority. In ordinary conversation, however, it is true that unqualified assertion (without modals, and without reported evidence) signals the strongest veridical commitment, and that assertions with reported evidence tend to trigger some doubt with the use of subjective expressions such as modals, attitude verbs (*seems, see, hear*, and the like).

This paper is, to our knowledge, among the first attempts to study the relation between veridicality judgment and evidence in online discourses. It is also probably the first study to ask what evidential categories are at play in those networks, and if these categories differ from those observed in ordinary co-operative conversation and in cross-linguistically attested grammatical systems. Strikingly, in contexts where pictures can be shared such as in Twitter, evidentiality is no longer strictly a grammatical category, but extends to encompass a variety of semiotic objects (pictures, videos, GIFs, link to news-media article etc.) that serve the purpose of grounding the veridicality judgment.

Our work addresses two specific questions. What are the evidential categories at play in tweets? And second, how are those evidential categories related to assertion and subjective statements? Our main finding is that assertions empirically correlate with indirect evidence in Twitter. To interpret this association, we propose that a speaker evaluates evidence as adequate (or even best possible)

grounds for veridical commitment because they deem it trustworthy. Trustworthiness is therefore the result of what we call a *trustworthiness evaluation*, which is a subjective measure that assigns a trustworthiness value to each piece of evidence no matter its type.

The discussion proceeds as follows. Section 2 reviews previous research on evidence types and assertion, with particular attention to reported evidence and testimony. We engage with linguistic, computational, and philosophical perspectives. In Section 3 we present our corpus study. We unfold the speech act annotation, our new annotation scheme of evidence type categories on Twitter, and we study the correlations between assertions and evidential basis. Section 4 provides a discussion of the main findings and presents our theoretical proposal highlighting the role of trustworthiness evaluation in shaping evidence type strategies. We argue that the type of assertive statements (bare or subjective) is a cue to determine the prior trustworthiness evaluation made by the speaker on the asserted content and on the evidence selected to ground this latter. Section 5 concludes.

2. Evidence and assertions: The existing picture

2.1 Evidence-type and evidential categories

Evidence has been extensively examined across three main fields of research: linguistic studies on evidentiality, NLP research on eyewitness accounts in emergency situations, and philosophical studies on the epistemology of testimony. In this section, we primarily focus on the first two areas, reviewing the different evidence typologies identified in these domains. The final subsection will address more in-depth the philosophical literature on the epistemology of testimony focus on the hapax of reported evidence.

2.1.1 Evidentiality across languages

Linguistic theorizing on evidentiality has produced very elaborate descriptions across languages and frameworks and overwhelmingly targets grammatical systems that encode the source of information.² In the literature, evidence is generally defined as the *justification* or *ground for a given speech act* (see a.o. Faller

2. It would be impossible enterprise to cite the voluminous literature; some foundational works are Chafe and Nichols (1986); Willett (1988); Nuyts (1992); De Haan (1999); Nuyts (2001); De Haan (2001); Garrett (2001); Faller (2002); Squartini (2004, 2012); Aikhenvald (2004, 2012); Matthewson et al. (2007); Matthewson *et al.* (2010); Speas (2008); Murray (2010); Faller (2011); Smirnova (2012); McCready (2015).

2002; Aikhenvald 2004; Murray 2010; McCready 2015; Murray 2017). In an assertive statement the ground is the information source on which the speaker based his own assertion.

Focusing the correlation between evidentiality and commitment, we focus on the quasi-universal distinctions identified in (Willett 1988; Aikhenvald 2004, 2018), specifically on the most important divide between DIRECT evidence (i.e., information acquired by the speaker witnessing the event) and INDIRECT evidence (i.e. information reported from a secondhand or a thirdhand source) (Faller 2002). Aikhenvald (2004) states, for example, that the Shipibo-Konibo language (a Panoan language spoken in the Peruvian Amazon) has an evidential system with four categories: Direct (or Visual), Inferred, Assumed, Reported. If a speaker is referring to something she has seen (Visual), she will add the clitic *ra* (see (2)).

- (2) *westíora nete-n- ra ka-a iki nokon yosi betan e -a, piti*
 One day-temp -DIR.EV go-PP2 AUX POSS1 elder CONJ 1-ABS fish:ABS
bená -i...
 search -SSSI
 ‘One day my grandfather and I went to look for fish (I see it)’

On the other hand, if the speaker is referring to an element he has not seen, but infers from direct evidence (**Inferred**), she will add the clitic *bira* (see (3)).

- (3) *mi-n bake pi-kas-bira-[a]i, oin-we!*
 2-GEN child:ABS eat-DES-INFR-INC see-IMP
 ‘Your child must be hungry (**inferred on the basis of him crying**), come and see!’

If the speaker is referring to an event she has not seen but infers it from general knowledge **Assumption** (or speculation), the clitic *-mein* will be added (see (4)). For example, if someone knocks on the door, one would use *-mein* to ask oneself ‘who could this be?’

- (4) *tso- a -mein i-ti iki*
 who- ABS -SPECL be-INF AUX
 ‘Who could it be?’ (speculation from general knowledge)

If a speaker is reporting the words of a general hearsay or quoting someone directly (**Reported**), he will use the clitic *ronki* (see (5)).

- (5) *a- ronki -a iki*
 do.TR- REP -PP2 AUX
 ‘It is said that s/he did (it)’ or ‘S/he says that s/he did (it)’

The cross-linguistic typology established by Aikhenvald (2004, 2018) identifies six evidential categories, that is to say, six types of information source linguistically marked in the languages studied as shown in Table 1. Her classification includes **Visual**, **Non visual sensory**, **Inference**, **Assumption**, **Hearsay** and **Quotative**.

Table 1. Recurrent semantic parameters in grammatical evidentiality (Aikhenvald 2004, 2018)

Visual	It covers information acquired through seeing
Non visual sensory	It covers information acquired through hearing, is typically extended to smell and taste, and sometimes also touch
Inference	Reasoning based on visible or tangible results
Assumption	Inference reasoning based on information other than visible results this may include logical reasoning, assumption, or simply general knowledge
Reported	It covers the reported information with no reference to who it was reported by
Quotative	It covers the reported information with an overt reference to the authorship of the quoted source

In a similar vein, Willett (1988) proposes a typology of evidentials organized around the distinction between Direct and Indirect sources, with the latter being further elaborated into the three sub-categories of **Reported** (i.e., information from a second or third-hand source such as a specific person, a rumor, popular folklore, etc.), **Inference** (i.e., inference or hypothesis based on visible or mental evidence), and **Attested** (i.e., first-hand information related to the senses (sight, hearing, touch, etc.)) as represented in Figure 1.

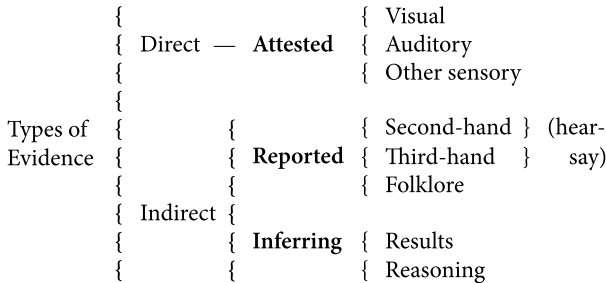


Figure 1. Classification of evidentiality Willett (1988)

One of the aims of our study is to investigate whether these variants are observed in social media.

2.1.2 Evidence-type in the literature on social media

Attention to evidence type has also been devoted in NLP (Natural Language Processing) research, particularly in emergency situations, where discriminating direct from indirect witness is critical. This strand of research examined source types, information credibility, and eyewitness categories to optimize information filtering in emergency situations.

Several classifications of source types (Olteanu *et al.* 2015) or witness types (Zahra *et al.* 2020) have been carried out. Zahra *et al.* (2020) proposed a classification of witness types. They identified five categories of eyewitnesses: **Direct**, **Indirect**, **Vulnerable**, **No Testimony**, **Irrelevant** (see Table 2).

Table 2. Eyewitness types in urgency situation on Twitter in Zahra *et al.* (2020)

Eyewitness types	Definition
Direct eyewitness	Testimonies associated with the direct perception of an event through senses (hearing, seeing, touching, etc.). They relay first-hand information.
Indirect eyewitness	This category covers testimonies relayed by direct witnesses received from their family, or from a restricted social circle.
Vulnerable eyewitness	This category covers tweets in which the author shares advice, orders and alerts received from local authorities. The aim of this category is to detect users who have anticipated an emergency situation or who are likely to be directly involved in the crisis situation.
No testimony	Do not have any explicit characteristics of testimonials.
Irrelevant	Messages that are too ambiguous to be classified, but which may have the characteristics of testimonials.

Similarly, Olteanu *et al.* (2015) identified six main types of sources: **Witnesses** (for first-hand accounts from people directly affected by the crisis), **Government** (for information relayed by institutional or administrative authorities), **NGOS** (for information relayed by non-profit associations), **Companies** (for information relayed by profit companies or associations), **Media** (for information relayed from traditional or alternative media sources i.e. blogs or online forums), **Outsiders** (for accounts from people not directly affected by the crisis or second-hand accounts).

Beyond source classifications, NLP researchers have increasingly focused on assessing the credibility of information related to source types, a matter to which we now turn.

2.2 On the relation between evidentiality and assertivity

As noted earlier, assertion is associated with a veridical sincerity condition that the asserted propositional content is believed or known to be true. The strength of commitment is governed by norms: the norm of belief (a speaker must believe in the asserted content, see Davidson 1985; Vanderveken 1990; Lauer 2013, a.o.), the Knowledge Norm Account (a speaker must know the asserted content, Williamson (1996)), or the truth norm (asserting something is recognizing it as true (see i.a. with its maxim of quality Grice 1975; Dummett 1981; Wright 1994; Williams 2010; Giannakidou and Mari 2021a). Additionally, social norms also play a crucial role in the strength of the assertion imposing assertor's responsibility (or 'liability to sanctions' for Peirce 1974) and necessitating for the assertor to justify their statement when questioned (Brandom 1983) (for a review of these social conventions see Shapiro 2020).

Commitment to truth is a gradable notion, or as Liu (2019) characterizes it, *elastic*. In a prior literature, Vanderveken (1990), Giannakidou and Mari (2016a, 2021c) have highlighted different degrees of commitment, sometimes implied by terms like *suggest* and *guess* insofar as they involve different sincerity regarding the speaker's intentions.³ The strong commitment of the speaker in their assertion has often been contrasted with weaker commitment in modal sentences (Karttunen 1972; Giannakidou 2013; Giannakidou and Mari 2021c), polar questions (Farkas and Bruce 2010; Krifka 2015; Farkas and Roelofsen 2017), or belief-based statements (Hawthorne *et al.* 2016).

Unlike bare assertions, assertions with a possibility modal (*John might be home*) are inquisitive (Giannakidou 2013), and designate lower degree of commitment which leads to lower informativity. Veridical commitment can be envisioned as a scale that distinguishes between full, partial, and trivial commitment: (6) (Giannakidou and Mari 2016b, 2021c).

- (6) Scale of veridical commitment:
 $\ll p, \text{MUST } p, \text{MIGHT } p \gg$;

3. Gradability of commitment also extends in the literature on modal adverbs (Yalcin 2007; Portner 2009; Lassiter 2010) building on scale structures (Kennedy and McNally 2005).

where i is the speaker, p conveys *full* commitment of i to p ; **MUST** p conveys *partial* commitment of i to p , and **MIGHT** p conveys *trivial* commitment of i to p . The order is from the strongest to the weakest expression of commitment.

The selection of information impacts the degree of commitment. Experiments have shown (Robinson *et al.* 1995) that speakers in conversation have the cognitive capacity to prioritize information from the most reliable sources, distinguishing information from reliable sources from information from unreliable sources. Within this framework, evidence that is proximally accessible to the speaker (such as direct perception) is typically more reliable than mediated or reported evidence (indirect evidence) (see for a survey of this philosophical tradition among others Evans 1982; Coady 1992; Shapin 1995; Lackey 2006; McCready 2015; Goldberg 2020). Hence, data communicated through a direct source tends to elicit a heightened commitment level leading to greater credibility in contrast to information conveyed through an indirect source (see Figure 2). This correlation had been highlighted in the semantic and typological literature by De Haan (1999). Each information source is considered to lay on a scale of credibility influencing the speaker's veridical commitment. The idea is supported by Frajzyngier (1985); Palmer (1986) that 'different manners of acquiring knowledge correspond to different degrees of certainty about the truth of the proposition' (Frajzyngier 1985, cited by De Haan 1999). The indirect source holds a particular status (as we will see more in depth in the next subsection): it indicates veridical disengagement from the speaker since another source is explicitly mentioned, and the information it relays can be denied by the speaker without leading to a contradiction (see *i.a.* Karttunen 1972; Faller 2002, for an analysis of weak sources as indirect; see Von Fintel and Gillies 2010; AnderBois 2014, for another view that considers indirect sources as not weak). Saurí and Pustejovsky (2009) proposes a computational analysis resonating with the same picture based on factuality in relation to information sources. A direct source would commit the speaker more strongly on the factuality of the event described than an indirect source.

This association between credibility and the type of information source has also been observed in the NLP literature on social network analysis in the context of automatically differentiating between credible and non-credible information

Evidential hierarchy

visual < auditory < nonvisual	<	inference ⁸ < quotative
direct evidence	<	indirect evidence

more believable < ----- > less believable

Figure 2. Relation between evidentiality and credibility in De Haan (1999)

sources (for an extensive review of the literature on credibility, see Castillo *et al.* (2011)). Both the type of sources disseminated on platforms like Twitter (Olteanu *et al.* 2015) and the type of eyewitnesses (Zahra *et al.* 2020) have been identified as crucial markers for discerning credible information (i.e. authentic) from non-credible information (i.e. misinformation or fake news). We will spell out the association between evidence type and credibility at the end of our paper in the discussion part. But let us first present some thoughts on the different epistemological debates that surround reported evidence and testimony.

2.2.1 *Reported evidence: An unreliable evidence?*

As discussed in the previous subsection, different types of evidence have been attributed distinct inherent epistemic values in various fields of research. It has been widely recognized that direct evidence is more likely to be perceived as conveying truthful content, whereas indirect evidence is often regarded with skepticism.

Within this epistemological framework, reported evidence – i.e., second-hand or third-hand information – is frequently classified as unreliable due to its increased susceptibility to distortion or misrepresentation. The inherent skepticism toward testimonial evidence aligns with what Coady (1992) terms ‘negative epistemology,’ a philosophical tradition that questions the reliability of testimony and denies it the status of knowledge or justified belief.

Philosophical skepticism toward testimonial evidence, or information acquired from the testimony of others, can be traced to two main traditions in epistemology:

Reductionist View: Testimony alone does not constitute knowledge; it must be corroborated by external evidence (or facts) to be qualified as justified belief or knowledge. This perspective is rooted in the empiricist tradition, particularly in Hume’s *An Inquiry concerning Human Understanding*, who emphasized habit and experience as justifications (habit to observe a match between facts and testimony). Similarly Mackie (1969) argues that individual observation is necessary for knowledge acquisition, contrasting with the opposing stance found in Reid (1850), which defends testimony as a fundamental source of knowledge.

Conservative View: Reported evidence and testimony are less likely to yield truthful content because the speaker lacks direct epistemic access to the information. The multiple layers of mediation introduce a risk of distortion. This position is exemplified in Locke’s *An Essay Concerning Human Understanding*, as referenced by Coady (1992) but also a.o. Lackey (2006)).

This skepticism toward reported evidence extends to legal philosophy. In their *Analysis of Evidence*, Anderson *et al.* (2005) recognize that different types of evidence hold distinct epistemic status concerning their ability to serve as proof in

court. In English common law, testimony must generally be firsthand, and hearsay evidence (i.e., reported evidence) is deemed inadmissible in most cases.⁴

Beyond legal contexts, skepticism toward reported evidence is embedded within a broader tradition of cognitive autonomy, which Coady (1992) terms the ‘individualist ideology.’ This view posits that only individual reasoning or direct observation can lead to knowledge and justified belief. This notion can be traced back to Plato’s *Theaetetus* (201d), where it is argued that testimony in trials does not constitute knowledge, even if it can guide decisions. Similarly, Descartes, in *Discourse on the Method*, adheres to negative epistemology, asserting that ‘a single man is much more likely to hit upon [proofs of truth] than a group of people.’ This tradition of epistemic egoism (a term we borrowed from Foley (1994)) has been extensively analyzed by Coady (1992); Lackey (2006); Adler (2006); Origgi (2004, 2006), and others. However, such a perspective arguably misrepresents scientific knowledge, which fundamentally relies on collective labor and a division of expertise (see Weber 1919; Kuhn 1962; Merton 1973; Latour and Woolgar 1979; Bourdieu 2001; Goldberg 2011) and daily-life knowledge acquisition as David Hume *An Inquiry concerning human understanding* (5, Section 10) highlighted ‘there is no species of reasoning more common, more useful, or even more necessary to human life, than that which is derived from the testimony of men and the reports of eyewitnesses and spectators.’

More recently, linguistic studies have revisited these debates by examining the interaction between veridical commitment and reported evidence. It has been argued that reported evidence may attenuate the illocutionary force of assertive statements, thereby expressing a weaker veridical commitment of the speaker to the asserted content. This perspective aligns with the *Weak Assertion Framework* and we will present them next.

2.2.2 *Weaker assertion?*

The possibility of asserting with an indirect source raises the question of whether the assertion itself becomes weaker. One response is to indeed weaken the concept of assertion itself so that we have stronger and weaker assertions (the gradual theories of assertion). The second response, which we shall adopt in our study, is to acknowledge that veridical commitment remains robust but we must allow for the possibility that the threshold for what counts as adequate or reliable evidence for commitment can be lowered; and this is so because assertions in Twitter have different goals than assertions in ordinary conversations (for the different goals of assertions on Twitter see Boscaro *et al.* (2024a,b)).

4. Other legal systems operate differently and jurisdictions may accept reported testimony based on the principle that a witness can transfer their epistemic authority to others through justifying reported facts.

Within the gradual theory of assertions, evidentials are treated as being able to modify the illocutionary force thus resulting in strong and weak assertion types (for a comprehensive analysis on weak assertion see Mandelkern 2022). Speech acts, including assertions, are subject to both syntactic and, notably, compositional modifications (Cohen and Krifka 2014; Greenberg and Wolf 2018;⁵ and more specifically for assertion, Greenberg and Wolf 2018; Incurvati and Schlöder 2019; Krifka 2023). Incurvati and Schlöder (2019) analyze the adverb *perhaps* as a modifier weakening the assertion's illocutionary force. Greenberg and Wolf (2018) analyze assertions as inherently gradable entities, denoting a modifiable degree of credence by both covert and overt modifiers. Modal adverbials (*possibly, probably, necessarily*) are deemed overt modifiers of assertion force altering the speaker's degree of commitment to the asserted content.

Similar to modal adverbials, evidentials exert a comparable influence on assertions, according to Krifka (2023) who suggests that the employment of the indirect evidential *apparently* should be construed as weakening the assertive force of the statement. Meanwhile, Faller (2002) proposes interpreting the reportative evidential *si* in Cusco Quechua as weakening the illocutionary strength of the assertion, transforming it into the alternative speech act called *Presentative*: the use of *si* substitutes the sincerity conditions of the assertion (i.e., the speaker's belief in the content uttered) with an alternative sincerity condition (i.e., the speaker's presentation of evidence for the content uttered).

This alteration in the force of an assertion (either weakening or strengthening) has also been elucidated through social and conversational perspectives. For instance, Krifka (2023) examines the attenuation of assertive force as a strategy to mitigate its social ramifications and remaining credible to interlocutors. A 'strong' assertion, if disproven, entails a greater social cost compared to a 'weak' assertion. Incurvati and Schlöder (2019), on the other hand, scrutinize the adjustment of assertive force as a mechanism to prevent a categorical negation by the interlocutor or the inclusion of an assertion into the common ground when the asserted proposition lacks adequate reliability.

Davis *et al.* (2007) posit that evidentials alter the 'quality threshold' of the contextual utterance (Potts 2007) and substitute it with a new 'quality threshold'. In the scenario of an assertion featuring a reported source, the quality threshold is lower compared to that of an assertion with a direct source, implying that the context is, in this respect, more tolerant with reduced veridical commitment on the part of the speaker.

5. Prior literature has introduced other modifiers of illocutionary forces for speech acts (Holmes (1984) for a comprehensive list, Sbisà (2001), Wolf (2015), Yatsushiro *et al.* (2022) for subjective epistemics).

As we now show, the association between relayed evidence and assertion is not an exception or an unexpected phenomenon in social networks. It is conceivable and foreseeable that social networks may favor indirect sources of information –indeed they are defined as platforms of information sharing (Kwak *et al.* 2010). What this default association shows is that assertion is not weakened, but that it serves purposes other than merely adding information to the common ground (see also Boscaro *et al.* 2024a, b). We proceed by presenting the corpus study in Section 3 and interpret the results in Section 4.

3. Evidence and veridical commitment in social media

For this study, we rely on a corpus of > 19K French tweets, originally collected by Kozłowski *et al.* (2020),⁶ and augmented by Bourgon *et al.* (2022) that gathers messages related to sudden crises (flooding, explosions, fires, etc.), and annotated according to the speech-act annotation schema in Laurenti *et al.* (2022a, b).⁷

We have added the annotation for evidence type categories. We present these two annotations in Sections 3.1 and 3.2 and presents the main results pertaining to the relevant correlations between speech acts categories and evidence type categories in Section 3.3.

3.1 Speech act categories

Discovered by Austin (1962) and widely promoted by Searle (1975), speech acts have been the subject of much discussion in philosophical and linguistic literature (Hamblin 1970; Brandom 1994; Sadock 2004; Asher and Lascarides 2008; Portner 2018, to name but a few). In Austin's original vision, LAs are intended to perform an action rather than convey information. When the priest pronounces the words '*I now pronounce you man and wife*', he is performing the action of marrying people rather than simply stating a proposition.

Beyond these prototypical cases, the literature has rapidly expanded understanding of the notion of speech acts as a particular type of linguistic object encompassing questions, commands and assertions, and going beyond proposi-

6. https://github.com/DiegoKoz/french_ecological_crisis

7. The corpus gathers messages related to ecological crises that occurred in France from 2016 to 2022 and posted 24h before, during (48h) and up to 72h after the crisis: 2 floods that occurred in Aude and Corsica regions, 8 storms (Béryl, Berguitta, Fionn, Eleanor, Bruno, Egon, Ulrika, Susanna), 2 hurricanes (Irma and Harvey), 2 building collapses (Marseille, Lille), 2 chemical plants explosions (Lubrizol, Sanary), 2 fires (Notre-Dame fire, Gironde and Landes wildfires). The data comes with additional metadata including: number of likes, retweets, followers and followings of the user.

tional content to reveal the communicative intentions of the speaker (Bach and Harnish 1984; Gunlogson 2008; Asher and Lascarides 2008; Farkas and Bruce 2010; Ginzburg 2012; Krifka 2015; Portner 2018, to cite just a few). Speech acts can in fact also be understood as *attitudes* or *stances* towards propositional content p in line with Giannakidou and Mari (2021c): by *asserting*, the speaker presents p as true, and as worth being added to the common ground; by *questioning*, she reveals her uncertainty towards p ; by *ordering*, she requests that p be realized; and by *exclaiming*, the speaker expresses a subjective evaluation of p .

Speech acts have been the subject of an extensive body of work in the computational linguistics literature and have proven to be an important preliminary step in several NLP applications, such as dialogue summarization Goo and Chen (2018) and conversational systems Higashinaka *et al.* (2014). Laurenti *et al.* (2022a, b) propose an annotation schema both at the tweet and the sub-tweet level focusing on natural disasters. Their classification envisions speech acts as stances towards p , their research is therefore of particular relevance for us, insofar as it considers tweets that are about a well identifiable event (storms, floods, ...). These events are the fact of the matter under discussion and about which truth is stated, evaluations are issued, action is required, and information is asked. The main four-speech act categories at the tweet level are indeed the following ones:

- **Jussives.** The tweets of this category enhance commitment to take action. Laurenti *et al.* (2022a, b) distinguish the following types of jussives: *commissives* (the speaker commits himself or herself to action), *exhortative* (the speaker commits some relevant individuals to action), *orders* (the speaker commits the addressee to taking action), and *open options* (the speaker describes the existence of a possibility of taking action).
 - (7) Va falloir ancrer le fauteuil au sol @philousports #Eleanor
<https://x.com/On3Fr3nch4T53P/status/948215772677050369>
 ‘We’ll have to anchor the chair to the ground @philousports #Eleanor’
<https://x.com/On3Fr3nch4T53P/status/948215772677050369>
- **Assertives.** Assertions are declarative sentences that convey objective truth (in opposition to subjective truth, in subjectives, cf. *infra*).
 - (8) Immeuble effondré à Sanary-sur-Mer dans le Var: un homme retrouvé mort https://www.bfmtv.com/var/effondrement-d-un-immeuble-a-sanary-sur-mer-le-difficile-travail-des-pompiers-pour-retrouver-les-disparus_AV-202112070434.html
 ‘Collapsed building in Sanary-sur-Mer, Var: a man found dead’ https://www.bfmtv.com/var/effondrement-d-un-immeuble-a-sanary-sur-mer-le-difficile-travail-des-pompiers-pour-retrouver-les-disparus_AV-202112070434.html

- **Interrogatives.** This category generally refers to tweets in the interrogative form that seek information.⁸
 - (9) Est-ce que la responsabilité de la ville de Lille est engagée?
'Is the city of Lille liable?'
- **Subjectives.** This category refers to tweets with a declarative form (like Assertive statements). The speaker shares a mental state in these tweets, which may be an *evaluation*, a *preference*, or an *expressive state* (emotion, feeling). In the literature on evidentiality, subjectivity is often considered an evidential category that encompasses information based on emotions, credence and personal preferences (Nuyts 1992; Garrett 2001; Squartini 2010; Grano 2021; Giannakidou and Mari 2021a). Here we consider subjectives as revealing a specific form of commitment. By using a subjective, the speaker signals that they are not ready to add the informational content to the common ground (Farkas and Bruce 2010; Krifka 2015) and we hereby consider subjectivity as a modal notion rather than an evidential one.
 - (10) Le bébé sauvé dans l'explosion à sanary c'est vraiment un miracle
'The baby rescued in the explosion in sanary is truly a miracle'

Finally Laurenti *et al.* (2022a, b) added to their classification the category **Others** which corresponds to the tweets where the speech act is difficult to identify or is uncertain.

- (11) Le bruit et l'odeur #Chirac https://www.youtube.com/watch?v=Fh2otDHTo68&ab_channel=FredLille
'The noise and the smell #Chirac https://www.youtube.com/watch?v=Fh2otDHTo68&ab_channel=FredLille'

Table 3 shows the distribution of speech acts in the corpus of 3137 tweets spanning over 7 crises, which we have selected to perform our double annotation for speech acts and evidence type. Assertives are the large majority, immediately followed by subjectives.

With this speech act distribution in our corpus, we now turn to the annotation schema for evidence categories in social media.

3.2 Evidence type categories and annotation schema

The Annotation was performed by two independent annotators with a kappa Cohen of 0.7.

8. Biased questions that reveal a preference of the speaker are Subjectives.

Table 3. Speech acts Annotation Distribution

Speech acts	Count	%
Assertive	1892	60.31%
Subjective	688	21.93%
Jussive	461	14.70%
Interrogative	96	3.06%
Total	3137	100.00%

Our study of evidence types on social networks relies on the application of the typologies of Aikhenvald (2004, 2018) and Willett (1988). We found that the most common distinction (as seen in Section 2.1) is between **Direct** and **Indirect** evidentiality. However, the landscape of evidence is reshuffled in social media because the majority of evidence is marked extralinguistically with the use of pictures, hyperlinks, videos, etc. Evidence is not only introduced by linguistic marking, but it is also introduced by the extralinguistic elements we summarize in Table 4 below.

Table 4. Different forms of extralinguistic evidence on X

Evidence forms	Description
Hyperlinks	The hyperlink can redirect the reader to a news media article (most common), an online forum, a YouTube video, a video from another user account, a picture from another user account, a GIF, a post from another user, a drawing, a poster etc.
Mention @	The mention can be towards the account of a traditional news media, of a political agent, or of another user.
#	The # can take the form of #sourcename (such as #BFMTV) or it can be # related to some keywords related to the crisis, or some Twitter Trend.

Our corpus leads us to distinguish four main types of evidence. For each type we provide a definition and spell out the overlaps with more traditional categories, especially with the one identified in linguistic literature reviewed in subSection 2.1.

- **Direct**, as defined by Willett (1988), is the traditional first-hand witness category. The source may be Visual or, more generally, sensory (smell, touch, hearing). In the specific case of social media, pictures accompanying the message that are not relayed are considered to belong to the Direct category.

- (12) Notre Dame en feu. #feunotredame #notredame #paris pic.twitter.com/eSiCJ5WW6s
'Notre Dame in fire #firenotredame #notredame #paris pic.twitter.com/eSiCJ5WW6s'
- (13) Le feu grignote la toiture de Notre Dame #NotreDame pic.twitter.com/qasKOox3eS
'Fire consumes the roof of Notre Dame #NotreDame pic.twitter.com/qasKOox3eS'
- **Relayed.** This category overlaps with the *Reported* – third-hand source of Willett (1988), and the *Quotative* category of Aikhenvald (2004, 2018). This category encompasses cases featuring second or third hand information *explicitly stated in the tweet content*. In relayed tweets the information source is a media, a forum, or an online video. Relayed pictures, hyperlinks and mentions referring to a third-party source (media, institutional source, etc.) are considered Relayed.

(14) Immeubles effondrés à Lille: le gouvernement confie une mission d'étude à des experts pour déterminer les causes du drame <https://t.co/ID9anhScWv>
'Collapsed buildings in Lille: the government commissions experts to investigate the causes of the tragedy <https://t.co/ID9anhScWv>'

(15) Je connaissais cette famille: l'emotion du maire de Sanary après le dramatique effondrement d'un immeuble <https://t.co/LhPMwcJixh> <https://t.co/B6w4AlkCJq> (source VarMatin)
'I knew this family: the emotion of the mayor of Sanary after the dramatic collapse of a building <https://t.co/LhPMwcJixh> <https://t.co/B6w4AlkCJq> (source VarMatin)'

 - **Loose sources.** This category overlaps with the *Reported* category (Willett 1988) with a third-hand source or even the *Rumor's* category (Aikhenvald 2004, 2018). It can also correspond to Aikhenvald's *Hearsay* category, which is defined as a reported information with no reference to whom the source of that information is. The information in this category is difficult to identify but it is relevant and related to the annotated crisis.

(16) #VentViolent cette nuit: forte migration de Normands vers l'Alsace...
'#ViolentWind tonight: strong migration from Normandy to Alsace...'

 - **No Testimony.** This category is not an evidential category but has been added to our classification in order to identify tweets that are somehow related to the crisis at hand. It corresponds to the usual 'other' class.

- (17) marseille mdrrrrr
'Marseille lmaooooo'
- (18) Des formations de qualité #TransitionEnergétique, #batiment, #bioenergies, #territoires, #EnR, #economiecirculaire, #Industrie en 2018 et 2019 ? Découvrez le catalogue de #formation en ligne de @nWinstitut, membre du réseau @assoCLER #Datadock
'Quality training courses #Energytransition #building #bioenergies #territories, #EnR, #circulareconomy, #Industry in 2018 and 2019? Discover the online #training catalogue of @nWinstitut, member of @assoCLER network #Datadock'

Table 5 summarizes our classification for evidentiality in Twitter.

Table 5. Evidence type classification

Evidence type	Definition
Direct	Information with first-hand evidence (often a picture or video that has been taken by the speaker).
Relayed	Information with second-hand or third-hand evidence explicitly mentioned by a hyperlink, a hashtag # or mention @.
Loose evidence	Information with second-hand or third-hand evidence not mentioned in the tweet.
No evidence	There is no evidence for the information relayed, or the tweet is unrelated to the topic, which is not mentioned in the tweet.

3.3 Results: The assertive – relayed correlation

Table 6 shows the distribution of the annotation of evidence type. The most frequent category is **Relayed**, which accounts for 62.54% of the corpus, followed by **No Testimony** with 17.18%. The categories **Loose Sources** and **Direct** are less frequent, representing 13.23% and 7.04% of the corpus, respectively.

We have studied the correlations between speech act categories and evidence type categories and performed the Spearman tests on the 16 combinations. In the table we highlight the significant positive correlations and the slightly positive correlations.

Table 6. Evidentiality annotation distribution

Evidence types	Count	%
Direct	221	7.04%
Relayed	1962	62.54%
Loose sources	415	13.23%
No testimony	539	17.18%
Total	3137	100.00%

Table 7. Evidentiality vs speech acts

Evidentiality	Assertive	Subjective	Interrogative	Jussive	Total
Direct	123 (3.92%)	75 (2.39%)	6 (0.19%)	17 (0.54%)	221 (7.04%)
Relayed	1442 (45.97%)	161 (5.13%)	33 (1.05%)	326 (10.39%)	1962 (62.64%)
Loose sources	150 (4.78%)	217 (6.92%)	26 (0.83%)	22 (0.70%)	415 (13.23%)
No testimony	177 (5.64%)	235 (7.49%)	31 (0.99%)	96 (3.06%)	539 (17.18%)
Total	1892 (60.31%)	688 (21.93%)	96 (3.06%)	461 (14.70%)	3137 (100%)

We perform a χ^2 test of independence between. The χ^2 statistics takes a value of 717.54, which correspond to a p-value smaller than 10^{-10} . So we reject the null hypothesis that the phrase type and the phrase mood are independent, and conclude that there is a very strong relation between evidential sources and speech act type.

The table of relative distances to independence, is given below:

Table 8. Evidentiality vs speech acts: Relative difference to independence

Evidentiality	Assertive	Subjective	Interrogative	Jussive
Direct	-10.29	26.53	-0.76	-15.48
Relayed	258.67	-269.30	-27.04	37.67
Loose sources	-100.30	125.98	13.30	-38.99
No testimony	-148.08	116.79	14.51	16.79

Positive difference to independence indicates over-representation of a given evidential category in messages categorized for a given speech act category. Conversely negative difference to independence indicates an under-representation of a given evidential category in messages belonging to a given speech act category. Specifically, Relayed evidentiality is widely over-represented in Assertive sentences, and widely under-represented among Subjective sentences. All the same, Loose-sources evidentiality and No-testimony evidentiality are widely over-represented in Subjective messages.

Overall, setting aside Jussives and focusing on non-addressee oriented speech acts (Assertives and Subjectives),⁹ the data show that subjectivity is grounded in very uninformative sources or even lack of sources. This is expected insofar as subjective statements convey an internal mental state whose source is the belief or preferences of the speaker and do not necessarily need external evidence to be formed (see Nuyts 1992; Squartini 2012; Giannakidou and Mari 2021a). As for assertions, which are the main focus here, and which overwhelmingly rely in relayed information, we provide a discussion in the following section.

4. Evidence and trustworthiness evaluation

The results presented in the previous section, particularly the strong correlation between bare assertive statements and indirect evidence (as highlighted in Table 8) present an apparent puzzle for the relationship between veridical commitment and evidence when considered in light of the prior literature we started with, on reported evidence and testimony – as tweets seem to rely on reported evidence.

For the weak assertion hypothesis outlined in Section 2.2, the correlation may suggest that tweets are non-cooperative assertions: social media users do not appear to conform to traditional norms of assertion, particularly the principle of veridicality and conventional evidential scales:

(19) *Tweets as Uncooperative assertions* (attempt):

Social media users do not select appropriate evidence to support their strong veridical commitment. By choosing weak evidence such as reported evidence, they effectively weaken their veridical commitment.

Yet, we argue in this section, this is not the correct characterization of what is happening. Key to our argument is acknowledging that what counts as ‘adequate evidence’ for veridical commitment relies on trust – but trust does not rely auto-

9. Jussives and interrogatives are traditionally considered addressee-oriented.

matically on the objective hierarchies of evidence that privilege first-hand account (where the source can thus be omitted). Contrary to prior accounts, we argue that what renders evidence reliable for commitment on Twitter involves an inherently subjective *trustworthiness evaluation* which depends on the speaker assessing the source of evidence as trusted therefore as passing the threshold for veridical commitment. Mentioning the source is thus an asset for the evaluation and establishes that the source can be trusted.

The trustworthiness evaluation is a subjective metric that assigns a probability to the asserted content based on the type of evidence used and the speaker's appraisal of it. In other words, evidence is deemed adequate for trust insofar as the speaker considers it so. As a result, the notion of *best possible grounds* is necessarily subjective and shaped by the speaker's informational basis with some endogenous and exogenous reasons (see Section 2). In this discussion, we defend below the *Trustworthiness Principle* as an alternative to the weak assertion hypothesis:

- (20) *Trustworthiness Principle*: Trustworthiness is a subjective measure that assigns a probability to the content of an assertion based on the type of evidence presented, and the speaker's appreciation of that evidence. It is the prerequisite to veridical commitment.

The Trustworthiness Principle, as stated, completes the Veridicality Principle of assertion by stating the process that establishes veridical commitment. To articulate this proposal clearly we provide a brief overview of how trustworthiness has been conceptualized in philosophy and natural language processing – which informs our understanding of trustworthiness evaluation. We will argue that trustworthiness plays a crucial role in evidence selection and in determining whether a piece of evidence is perceived or not as the best possible evidence to ground one's veridical commitment.

4.1 Conceptions of trustworthiness

While relatively underexplored in linguistic research (except McCready 2015; Krifka 2023), trustworthiness has been extensively studied in natural language processing (NLP), philosophy, psychology, and also economics. It is distinct from related concepts such as trust, mere reliance, optimism, expectation, gullibility, and epistemic vigilance.

Trustworthiness in social media analysis Within NLP and social media analysis, trustworthiness is typically studied with credibility (for an overview see a.o. Fogg 2003; Fogg *et al.* 2003). Credibility has been related both to content and the source of a message (Wathen and Burkell 2002). Research in this domain has proposed various credibility metrics (Choi and Stvilia 2015) based on speaker

attributes such as honesty, neutrality, and expertise (Hovland and Weiss 1951). Trustworthiness, in this framework, is analyzed as one indicator of credibility assessment – the other indicator being expertise and competence (Hovland and Weiss 1951; Fogg 2003; Fogg *et al.* 2003). It has been defined as a *moral virtue* – i.e. perceived goodness and morality of the source, where the source appears as fair, unbiased, honest, and with integrity. Other external factors have been identified as indicators of this notion of trustworthiness including ethical behavior of the source, his commercial interests, his decency, his transparency.

Trustworthiness in Philosophy In the philosophical literature, trustworthiness is often conflated with trust but remains conceptually distinct: while trust is an attitude, trustworthiness is a property attributed to agents (Hardin 2002; Hieronymi 2008). This ability to assign trustworthiness has been acknowledged to arise very early before even the ability to discriminate between evidence types (Robinson *et al.* 1995; Robinson and Whitcombe 2003). It denotes an individual's likelihood of conveying truthful information, engaging in cooperative behavior – in the sense of Grice (1975) –, or fulfilling a commitment to an interlocutor. Several theoretical perspectives have been proposed:

Moral account (which is the most accepted one where Trustworthiness is viewed as a moral virtue (see Potter 2002).

Motivational account: It is assigned based on an agent's goodwill and cooperative disposition (see a.o., Baier 1986). Its cooperative disposition can also be motivated when both agents have same goal – as in the encapsulated motivations theory Hardin (2002).

Normative account: Trustworthiness is attributed to agents who adhere to specific norms (e.g., the principle of veridicality).

Institutional account: It is linked to institutional affiliation, which compels trustworthy behavior (Hardin 2002).

Affective account: Trustworthiness is treated as an affective attitude, akin to optimism or hope (Baier 1992).

Commitment account: Trustworthiness is assigned based on an individual's likelihood of fulfilling a given commitment (see a.o. Hardin 2002; Hawley 2014, 2019).

These perspectives highlight that trustworthiness is a property attributed to a person for diverse reasons – normative constraints, epistemic constraints, institutional affiliation, affective considerations, or commitment fulfillment. Additionally, research has emphasized variability in trustworthiness judgments, which depend on the speaker's role – e.g., *low trustee* vs. *high trustee* (Hardin 2002) – and on background assumptions that shape these evaluations (Hieronymi 2008).

4.2 Trustworthiness evaluation: View from social media

Our proposal challenges the assumptions found in philosophical epistemology and linguistic studies on evidentiality, which often posit that certain types of evidence inherently convey more truthful information than others. Instead, we argue that the evaluation of evidence is subjective and somewhat speaker-dependent: the same piece of evidence may be assessed differently depending on the speaker's subjective preferences, some background assumptions (as already claimed by Hieronymi (2008)), some facts that are directly or indirectly accessible in their physical environments, some norms of the conversation, some institutional constraints, etc. as we just discussed. Our perspective highlights the multidimensional nature of trustworthiness attribution as shown by the various debates on the definition of trustworthiness we saw.

We propose that there is no fundamental distinction between (21) and (22), as both cases involve assertions where the speaker expresses certainty towards, thus commitment to the asserted content – whether the evidence is direct or indirect.

(21) John went to the grocery store; he told me earlier.

(22) John went to the grocery store; I saw him there.

However, we maintain that there is a distinction between (23) and (24), in that although both involve indirect evidence one (24) is evaluated as certain by the speaker (with a strong veridical commitment), whereas the other (23) is assessed as inadequate leading to a weaker veridical commitment with modal bias, therefore only partial commitment (Giannakidou and Mari 2016b, 2018, 2021c, b).

(23) John must have been to the grocery store this morning; I saw him leaving earlier.

(24) John went to the grocery store this morning; I saw him leaving earlier.

Previous work, such as McCready (2015), has argued that evidential selection relies on a *function of reliability*, which ranks and evaluates evidence based on two factors: (i) an initial reliability assessment and (ii) a posterior evaluation based on the evidence's past accuracy. This account assumes that speakers adjust their assertions based on the perceived reliability of the evidence, particularly when the risk of relaying false information may impact their reputation (McCready 2015; Krifka 2023).

Rather than adopting this account that we consider to be externalist – as it supposes an acquaintance with both the evidence and its history of accuracy or inaccuracy –, we introduce a new framework in which both evidential selection and the evaluation of the conveyed content are explained through a trust-

worthiness evaluation. Unlike McCready (2015), we argue that knowledge of an evidence's past reliability is not a necessary condition for trustworthiness assessment. Instead, trustworthiness evaluation operates as a subjective measure, enabling speakers to assign either a positive value or a negative value independently of the evidence type (as stated in (25)).

- (25) *Trustworthiness evaluation (I)*: Trustworthiness is a subjective measure of an individual anchor i that assigns to the content provided by assertion a probability measure based on the evidence and i 's assessment of the evidence.

Speakers and hearers are individual anchors or judges. When selecting evidence and assessing the truthfulness of conveyed content, speakers consider multiple reasons. Matthewson (2010) identifies evidence type, evidence strength, and evidence location as necessary elements in the selection of evidence. We extend her analysis by proposing that evidential selection is based also on an evaluation of the speaker which is influenced by her informational basis, which consists of both endogenous (background assumptions, prior beliefs, subjective preferences, etc.) and exogenous (facts about the state of the world) components (as spelled out below in (26) and (27)) — see extensive discussion in Mari (2016); Giannakidou and Mari (2021a).

- (26) *Trustworthiness evaluation (II)*: The trustworthiness assigned to an assertion's content depends on the speaker's evaluation of both the evidence e and the proposition p it conveys. This evaluation is based on the speaker's informational basis, which includes (i) exogenous content C (factual information about the world) and (ii) endogenous content B (beliefs, preferences, and background knowledge of the speaker). A speaker considers e to be trustworthy if and only if p is supported by B and/or C .

The speaker's evaluation of trustworthiness thus results from these two components (C and B), leading to our following formalization of the trustworthiness evaluation function:

- (27) *Trustworthiness evaluation function (III)*: Let $\mu(\cdot \mid B \cap C)$ a contextually defined measure. A speaker A assigns to p the likelihood to be true based on B and C . A assigns p as trustworthy iff $\mu(p \mid B \cap C) = 1$.

To illustrate our proposal, we examine two scenarios (scenario 1 and scenario 2) drawn from our corpus observation, each presenting different mechanisms of trustworthiness evaluation. In Scenario 1 we present the trustworthiness evaluation which is at stake in Assertive statement grounded in *Relayed* evidence (where evaluation is based on both B and C). In scenario 2, we present the trustworthiness evaluation at stake in Assertive statements grounded in *Loose* evidence (where evaluation primarily relies on B).

Scenario 1: Assertions with Relayed evidence

We consider Relayed evidence **rp** in this scenario and its correlation with bare assertive statement, which is the most important correlation observed in our corpus as highlighted in Section 3.

(28) Assertions with relayed evidence.

- a. The speaker *A* relies on a source **rp** in context *c*. **rp** is the source of content *C*, and is either a hyperlink, or a mention @ or a #sourcename.
- b. Let *C* and *B* be the two components of the informational basis of *A*. *C* is the exogenous evidence, i.e., the content provided by the source **rp**. *B* is the set of subjective preferences and/or privately held beliefs of *A*. The source is trustworthy iff $\mu(p \mid B \cap C) = 1$.
- c. If the source is trustworthy and *p* is assertable: $\text{ASSERT}(A, p)$

Scenario 2: Assertions with Loose evidence

In this scenario, we consider Loose evidence (**le**) and its correlation with assertive statement. Assertions based on **le**, proceed from a different principle.

(29) Assertions with loose sources.

- a. The speaker *A* relies on a source **le** in a context *c*. **le** is the source which is not explicitly mentioned and with no content *C* directly provided. In this scenario, only one # merely determines the *QUD* – i.e. the question under discussion as defined by Roberts (1996); Ginzburg (1996, 2012).
- b. Let *B* and # be the two components of the informational basis *A*. # merely provides the *QUD* for *p*, while *B* is the set of subjective preferences and/or privately held beliefs of *A*. The source is trustworthy and is hereby assertable by *A* iff $p \subset B$.
- c. If the source is trustworthy and *p* is assertable: $\text{ASSERT}(A, p)$

As illustrated by Scenario 1 with **rp** evidence and by Scenario 2 with **le** evidence, one speaker might evaluate as trustworthy different types of evidence based on both factual content or/and subjective preferences, private beliefs, or background assumptions.

This subjective evaluation – the trustworthiness evaluation – is also signaled directly by the speech act uttered. While in producing a bare assertive statement, a speaker is strongly veridically committed to *p* and hereby evaluates the evidence and *p* as being trustworthy, in producing a modalized assertive statement (or as in our annotation schema a subjective statement, see Section 3), one speaker is weakly veridically committed to *p* and hereby evaluates the evidence and *p* as being less trustworthy. We consider that the type of speech act uttered by the speaker is therefore both a signal of its own veridical commitment (as already demonstrated by (6)), but also a signal of a trustworthiness evaluation function.

Our proposal highlights the fact that the type of speech act produced is not only a marker of the speaker's veridical commitment (as previously demonstrated in (6)) but also an overt signal of the trustworthiness evaluation function. Crucially, the speaker's commitment to a proposition is independent of the type of evidence grounding it but is influenced by the trustworthiness assigned to both the evidence and the propositional content. We argue that this evaluation process – formally captured by our trustworthiness evaluation function – is fundamental to understanding evidential selection and assertion strategies in social media discourse.



5. Conclusion

In conclusion, this paper provided a first analysis of evidence selection in Twitter and its relationship with the speaker's veridical commitment. We observed that speakers ground their assertive statements in various types of evidence, which encompass the traditional distinction between Direct and Indirect evidence established in prior linguistic literature.


Our statistical results indicate a strong correlation between Reported evidence and bare Assertions, suggesting that speakers often rely on reported information while still expressing a strong veridical commitment. We demonstrated that this should not be interpreted as noncooperative behavior, even if indirect evidence is typically considered inadequate for a strong veridical commitment. We argued that this perspective is too restrictive and that evidence type should be decoupled from reliability.

Instead, we propose that the perceived adequacy of an evidence type stems from a subjective evaluation of trustworthiness of the speaker of both the evidence and the propositional content it conveys. This evaluation is shaped by the speaker's informational basis, which includes subjective preferences, prior beliefs, background assumptions and biases, as well as external knowledge about the world.

References

- Adler, Jonathan E. 2006. "Epistemological problems of testimony." In *Stanford Encyclopedia of Philosophy*.
-  Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.
-  Aikhenvald, Alexandra Y. 2018. *The Oxford handbook of evidentiality*. Oxford: Oxford University Press.

- doi AnderBois, Sscott. 2014. "On the exceptional status of reportative evidentials." *Semantics and Linguistic Theory* 24: 234–254.
- doi Anderson, Terence, Schum, David and Twining, William. 2005. *Analysis of evidence*. Cambridge: Cambridge University Press.
- Asher, Nicolas and Lascarides, Alex. 2008. "Commitments, beliefs and intentions in dialogue." In *Proceedings of the 12th Workshop on the Semantics and Pragmatics of Dialogue*, 29–36. Citeseer.
- Austin, John L. 1962. *How to do things with words*. Oxford: Oxford University Press.
- Bach, Kent and Harnish, Robert M. 1984. *Linguistic communication and speech acts*. 2nd edition, Cambridge / London: The MIT Press.
- doi Baier, Annette. 1986. "Trust and antitrust." *Ethics* 96 (2): 231–260.
- doi Baier, Annette. 1992. "Trusting people." *Philosophical Perspectives* 6: 137–153.
- doi Beyssade, Clarie and Marandin, Jean-Marie. 2009. "Commitment: une attitude dialogique." *Langue française* 2: 89–107.
- Boscaro, Marie, Giannakidou, Anastasia, Mari, Alda, and Tinarrage, Valentin. 2024. 'Assertions, cooperativity and evidence on X.' In: *Proceedings of the 28th Workshop on the Semantics and Pragmatics of Dialogue*.
- Boscaro, Marie, Giannakidou, Anastasia, Mari, Alda, and Tinarrage, Valentin. 2024. 'Cooperative Norms for Assertions and Conversational Grounding in X.' In: *Conversational Grounding in the Era of LLMs Workshop Booklet*, 18.
- Bourdieu, Pierre. 2001. *Science de la science et réflexivité*. Raisons d'Agir.
- Bourgon, Nils, Benamara, Farah, Mari, Alda, Moriceau, Véronique, Chevalier, Gaëtan, Leygue, Laurent. 2022. "Are sudden crises making me collapse? measuring transfer learning performances on urgency detection." In *19th International Conference on Information Systems for Crisis Response and Management, ISCRAM 2022, Tarbes, France, May 22-25, 2022*, R. Grace and H. Baharmand (eds.), 701.709. ISCRAM Digital Library.
- doi Brandom, Robert. 1983. "Asserting." *Noûs*: 637–650.
- Brandom, Robert. 1994. *Making it explicit: Reasoning, representing, and discursive commitment*. Cambridge: Harvard University Press.
- doi Castillo, Carlos, Mendoza, Marcela and Poblete, Barbara. 2011. "Information credibility on twitter." In *Proceedings of the 20th international conference on World wide web*, 675–684.
- Chafe, Wallace L. and Nichols, Johanna. 1986. *Evidentiality: The linguistic coding of epistemology*. Norwood: Ablex.
- doi Choi, Wonchan and Stvilia, Besiki. (2015). "Web credibility assessment: Conceptualization, operational- ization, variability, and models." *Journal of the Association for Information Science and Technology* 66 (12): 2399–2414.
- Coady, Cecile Anthony John. 1992. *Testimony: A philosophical study*. London: Clarendon Press.
- doi Cohen, Ariel and Krifka, Manfred. 2014. "Superlative quantifiers and meta-speech acts." *Linguistics and philosophy* 37: 41–90.
- doi Davidson, Donald. 1985. "Communication and convention." In *Dialogue: An interdisciplinary approach*: 11–26.
- doi Davis, Christopher, Potts, Christopher and Speas, Margaret. 2007. The pragmatic values of evidential sentences. *Semantics and Linguistic Theory* 17: 71–88.



- De Haan, Ferdinand. 1999. "Evidentiality and epistemic modality: setting boundaries." *Southwest Journal of Linguistics* 18 (1): 83–101.
-  De Haan, Ferdinand. 2001. "The place of inference within the evidential system." *International Journal of American Linguistics* 67 (2): 193–219.
- Dummett, Michale. 1981. *Frege: Philosophy of language*. Harvard: Harvard University Press.
- Evans, Gareth. 1982. *The Varieties of Reference*. London: Clarendon Press.
- Faller, Martina T. 2002. *Semantics and pragmatics of evidentials in Cuzco Quechua*. Stanford: Stanford University.
- Faller, Martina T. 2011. "A possible worlds semantics for cuzco quechua evidentials." In *Proceedings of SALT XX. Cornell University: CLC Publications*, 660–683.
-  Farkas, Donka F. and Bruce, Kim B. 2010. "On reacting to assertions and polar questions." *Journal of semantics* 27 (1): 81–118.
-  Farkas, Donka F. and Roelofsen, Floris. 2017. "Division of labor in the interpretation of declaratives and interrogatives." *Journal of Semantics* 34 (2): 237–289.
-  Fogg, B.J. (2003). "Prominence-interpretation theory: Explaining how people assess credibility online." In *CHI'03 extended abstracts on human factors in computing systems*, 722–723.
- Fogg, B.J., Soohoo, Cathy, Danielson, David R., Marable, Leslie, Stanford, Julianne and Tauber, Ellen R. 2003. "How do users evaluate the credibility of web sites? A study with over 2,500 participants." In *Proceedings of the 2003 conference on Designing for user experiences*, 1–15.
- Foley, Richard. (1994). "Egoism in epistemology." *Socializing Epistemology*, Rowman & Littlefield.
-  Frajzyngier, Zygmunt. 1985. "Truth and the indicative sentence." *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"* 9 (2): 243–254.
- Garrett, Edward John. (2001). *Evidentiality and assertion in Tibetan*. University of California, Los Angeles.
-  Geurts, Bart. 2019. "Communication as commitment sharing: speech acts, implicatures, common ground." *Theoretical linguistics* 45 (1–2): 1–30.
- Giannakidou, Anastasia. 2013. "Inquisitive assertions and nonveridicality." *The dynamic, inquisitive, and visionary life of φ A feestschrift for Jeroen Groenendijk*, Martin Stokhof and Frank Veltman, ed. by Maria Aloni, Michael Franke, F. Roelofsen, 115–126.
- Giannakidou, Anastasia and Mari, Alda. 2016a. "Emotive-factive and the puzzle of the subjunctive." *Proceeding of CLS* 51: 181–195.
-  Giannakidou, Anastasia and Mari, Alda. 2016b. "Epistemic future and epistemic must: Nonveridicality, evidence, and partial knowledge." In *Mood, Aspect, Modality Revisited. New answers to old questions*, J. Blaszcak et al. (eds.), 75–124. Chicago: The University of Chicago Press.
-  Giannakidou, Anastasia and Mari, Alda. 2018. "A unified analysis of the future as epistemic modality: The view from greek and italian." *Natural Language & Linguistic Theory* 36: 85–129.
-  Giannakidou, Anastasia and Mari, Alda. 2021a. "A linguistic framework for knowledge, belief, and veridicality judgment." *KNOW: A Journal on the Formation of Knowledge*.

- Giannakidou, Anastasia and Mari, Alda. 2021b. "Modalization and bias in questions." University of Chicago and Insitut Jean Nicod.
-  Giannakidou, Anastasia and Mari, Alda. 2021c. *Truth and veridicality in grammar and thought: Mood, modality, and propositional attitudes*. Chicago: The University of Chicago Press.
- Ginzburg, Jonathan. 1996. "Interrogatives: Questions, facts and dialogue." *The handbook of contemporary semantic theory*.
-  Ginzburg, Jonathan. 2012. *The interactive stance: Meaning for conversation*. Oxford: Oxford University Press.
-  Goldberg, Sandy. 2011. "The division of epistemic labor." *Episteme* 8 (1): 112–125.
-  Goldberg, Sanford C. 2020. "Trust and reliance 1." *The Routledge handbook of trust and philosophy*, 97–108.
-  Goo, Chih-Wen and Chen, Yun-Nung. 2018. "Abstractive dialogue summarization with sentence-gated modeling optimized by dialogue acts." In *Proceedings of 7th IEEE Workshop on Spoken Language Technology*.
-  Grano, Thomas. 2021. *Attitude reports*. Cambridge: Cambridge University Press.
- Greenberg, Yaël and Wolf, Lavi. 2018. "Gradable assertion speech acts." In *NELS 48: Proceedings of the Forty-Eighth Annual Meeting of the North East Linguistic Society*, Volume 1, 271–280.
-  Grice, H. Paul. 1975. "Logic and conversation." In *Syntax and semantics 3: Speech acts*, P-Cole et al. (eds.) 41–58.
- Gunlogson, Christine. 2001. *True to form: Rising and falling declaratives as questions in English*. University of California Santa Cruz.
-  Gunlogson, Christine. 2008. "A question of commitment." *Belgian Journal of Linguistics* 22 (1): 101–136.
- Hamblin, Charles L. 1970. "Fallacies." *Tijdschrift Voor Filosofie* 33 (1).
- Hardin, Russell. 2002. "Trust and trustworthiness." *Russell Sage Foundation*.
- Harnish, Robert M. 1994. "Mood, meaning and speech acts." In *Foundations of Speech Act Theory: Philosophical and Linguistic Perspectives*, S.L. Tsohatzidis (ed.), 407–459. London: Routledge.
-  Hawley, Katherine Jane. 2014. "Trust, distrust and commitment." *Noûs* 48 (1): 1–20.
-  Hawley, Katherine Jane. 2019. *How to be trustworthy*. Oxford: Oxford University Press.
-  Hawthorne, John, Rothschild, Daniel and Spectre, Levie. 2016. "Belief is weak." *Philosophical Studies* 173: 1393–1404.
-  Hieronymi, Pamela. 2008. "The reasons of trust." *Australasian Journal of Philosophy* 86 (2): 213–236.
- Higashinaka, Ryuichiro, Imamura, Kenji, Meguro, Toyomi, Miyazaki, Chiaki, Kobayashi, Nozomi, Sugiyama, Hiroaki, Hirano, Toru, Makino, Toshiro and Matsuo, Yoshihiro. 2014. "Towards an open-domain conversational system fully based on natural language processing." In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers*, 928–939. Dublin: Dublin City University and Association for Computational Linguistics.
-  Holmes, Janet. 1984. "Modifying illocutionary force." *Journal of Pragmatics* 8 (3): 345–365.

- doi Hovland, Carl I. and Weiss, Walter. 1951. "The influence of source credibility on communication effectiveness." *Public Opinion Quarterly* 15 (4): 635–650.
- doi Incurvati, Luca and Schlöder, Julian J. 2019. "Weak assertion." *The Philosophical Quarterly* 69 (277): 741–770.
- doi Karttunen, Lauri. 1972. "Possible and must." In *Syntax and Semantics Volume 1*, 1–20. Brill.
- doi Kennedy, Chris and McNally, Louise. 2005. "Scale structure, degree modification, and the semantics of gradable predicates." *Language*: 345–381.
- doi Kozłowski, Diego, Lannelongue, Elisa, Saudemont, Frédéric, Benamara, Farah, Mari, Alda, Moriceau, Véronique and Boumadane, Abdelmoumene. 2020. "A three-level classification of french tweets in ecological crises." *Information Processing & Management* 57 (5): 1–46.
- doi Kratzer, Angelika. 1991. "Modality." In *Semantics: An international handbook of contemporary research*, A. von Stechow and D. Wunderlich (eds.).
- doi Krifka, Manfred. 2015. "Bias in commitment space semantics: Declarative questions, negated questions, and question tags." In *Semantics and linguistic theory*, Volume 25, 328–345.
- Krifka Manfred. 2023. (Joint work with Fereshteh Modarresi) "Modifications of assertive commitments and their effect on trustworthiness." *ERC Advanced Grant Horizon 2020 787929 SPAGAD: Speech Acts in Grammar and Discourse Construction of Meaning Series*. Stanford: Stanford University.
- Kuhn, Thomas. S. 1962. *La structure des révolutions scientifiques*. Paris: Flammarion.
- doi Kwak, Haewoon, Lee, Changhyun, Park, Hosung and Moon, Sue. 2010. "What is twitter, a social network or a news media?" In *Proceedings of the 19th international conference on World wide web*, 591–600.
- doi Lackey, Jennifer. 2006. *The Epistemology of Testimony*. Oxford: Oxford University Press.
- doi Lassiter, Daniel. 2010. "Gradable epistemic modals, probability, and scale structure." In *Semantics and Linguistic Theory*, Volume 20, 197–215.
- Latour, Bruno and Woolgar, Steve. 1979. *La vie de laboratoire: La production des faits scientifiques*. Paris: La Découverte.
- Lauer, Svend. 2013. *Towards a Dynamic Pragmatics*. Ph. D. thesis, Stanford University.
- Laurenti, Enzo, Bourgon, Neils, Benamara, Farah, Mari, Alda, Moriceau, Véronique and Courgeon, Camille. 2022a. "Give me your intentions, I'll predict our actions: A two-level classification of speech acts for crisis management in social media." In: *Proceedings of the 13th Language Resources and Evaluation Conference (LREC 2022)*, 4333–4343. Marseille: European Language Resources Association.
- Laurenti, Enzo, Bourgon, Neils, Benamara, Farah, Mari, Alda, Moriceau, Véronique and Courgeon, Camille. 2022b. "Speech acts and communicative intentions for urgency detection." In: *Proceedings of the 11th Joint Conference on Lexical and Computational Semantics*, 289–298, Seattle / Washington: Association for Computational Linguistics.
- doi Liu, Mingya. 2019. The elastic nonveridicality property of indicative conditionals. *Linguistics Vanguard* 5.
- Mackie, J. L. 1969. "The possibility of innate knowledge." In *Proceedings of the Aristotelian Society*, 245–257.
- doi Mandelkern, Mathew A. and Dorst, Kevin. 2022. "Assertion is weak." *Philosophers' Imprint* 22 (n/a).

- doi Mari, Alda. 2016. "Assertability conditions of epistemic (and fictional) attitudes and mood variation." In *Semantics and Linguistic Theory*, Volume 26, 61–81.
- doi Matthewson, Lisa. 2010. "Cross-linguistic variation in modality systems: The role of mood." *Semantics and Pragmatics* 3: 1–74.
- doi Matthewson, L., Davis, Henry and Rullmann, Hotze. 2007. "Evidentials as epistemic modals: Evidence from st'át'imcets." *Linguistic variation yearbook* 7 (1): 201–254.
- McCready, Elin S. 2015. *Reliability in pragmatics*, Volume 4. Oxford Studies in Semantics and Pragmatics.
- Merton, Robert K. 1973. *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: The University of Chicago Press.
- Murray, Sarah E. 2010. Evidentiality and the structure of speech acts. Ph. D. thesis, Rutgers University-Graduate School-New Brunswick.
- doi Murray, Sarah E. 2017. *The semantics of evidentials*, Volume 9. Oxford: Oxford University Press.
- doi Nuyts, Jan. 1992. "Subjective vs. objective modality: What is the difference." *Layered structure and reference in a functional perspective*, 73–97.
- doi Nuyts, Jan. 2001. "Subjectivity as an evidential dimension in epistemic modal expressions." *Journal of Pragmatics* 33 (3): 383–400.
- doi Olteanu, Alexandra, Vieweg, Sarah and Castillo, Carlos. 2015. "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*, pp. 994–1009.
- doi Origgì, Gloria. 2004. "Is trust an epistemological notion?" *Episteme* 1 (1): 61–72.
- Origgì, Gloria. 2006. "L'ontogenesi della confidenza." *Raisons Pratiques* 2006.
- Palmer, Franck. R. 1986. *Mood and modality*. Cambridge: Cambridge University Press.
- Peirce, Charles S. 1974. *Collected papers of Charles Sanders Peirce*, Volume 1. Cambridge: Harvard University Press.
- doi Pietrandrea, Paola. 2005. *Epistemic Modality. Functional Properties and the Italian System*. Amsterdam: John Benjamins.
- Plato. 1942. *Théétète*. Translated in French and annotated by Léon Robin. In *Platon, Œuvres complètes II*, 142–192. Bibliothèque de la Pléiade, NRF.
- doi Portner, Paul H. 2009. *Modality*, Volume 1. Oxford: Oxford University Press.
- Portner, Paul H. 2018. *Mood*. Oxford: Oxford University Press.
- Portner, Paul H. and Partee, Barbara H. 2008. *Formal semantics: The essential readings*. John Wiley & Sons.
- doi Potter, Nancy N. 2002. *How Can I Be Trusted?: A Virtue Theory of Trustworthiness*. Rowman & Littlefield Publishers.
- doi Potts, Christopher. 2007. "Conversational implicatures via general pragmatic pressures." In *New Frontiers in Artificial Intelligence: JSAI 2006 Conference and Workshops, Tokyo, Japan, June 5-9 2006, Revised Selected Papers*, pp. 205–218. Springer.
- Reid, Thomas. 1850. *Essays on the intellectual powers of man*. J. Bartlett.
- Ricci, Claudia and Rossari, Corinne. 2018. "Commitment phenomena through the study of evidential markers in romance languages."
- Roberts, Craige. 1996. "Information structure in discourse: Towards an integrated formal theory of pragmatics." *Semantics and Pragmatics* 5 (6): 1–69.

- doi Robinson, Elizabeth J., Mitchell, Peter and Nye, Rebecca. 1995. "Young children's treating of utterances as unreliable sources of knowledge." *Journal of Child Language* 22 (3): 663–685.
- doi Robinson, Elizabeth J. and Whitcombe, E. 2003. "Children's suggestibility in relation to their understanding about sources of knowledge." *Child Development* 74 (1): 48–62.
- Sadock, Jerrold. 2004. "3 speech acts." *The handbook of pragmatics*, 53.
- doi Saurí, Roser and Pustejovsky, James. 2009. "Factbank: A corpus annotated with event factuality." *Language Resources and Evaluation* 43 (3): 227–268.
- doi Sbisà, Marina. 2001. "Illocutionary force and degrees of strength in language use." *Journal of Pragmatics* 33 (12): 1791–1814.
- doi Searle, John R. 1975. "Indirect speech acts." In *Speech acts*, 59–82. Leiden: Brill.
- Shapin, Steven. 1995. *A social history of truth: Civility and science in seventeenth-century England*. Chicago: The University of Chicago press.
- Shapiro, Lionel. 2020. "Commitment accounts of assertion." *The Oxford handbook of assertion*, 73–97.
- Smirnova, Anastasia. 2012. "Evidentiality in bulgarian: temporality, epistemic modality and information source." *Journal of Semantics*: 1–54.
- doi Speas, Peggy. 2008. "On the syntax and semantics of evidentials." *Language and Linguistics Compass* 2 (5): 940–965.
- doi Squartini, Mario. 2004. "Disentangling evidentiality and epistemic modality in romance." *Lingua* 114 (7): 873–895.
- doi Squartini, Mario. 2010. "Where mood, modality and illocution meet: the morphosyntax of romance conjectures." In *Modality and Mood in Romance, Modal interpretation, mood selection, and mood alternation*, M. Becker, E.M. Remberger (eds.). Tübingen: Niemeyer.
- doi Squartini, Mario. 2012. "Evidentiality in interaction: The concessive use of the italian future between grammar and discourse." *Journal of Pragmatics* 44 (15): 2116–2128.
- doi Stalnaker, Robert C. 1978. "Assertion." In *Syntax and Semantics 9: Pragmatics*, P. Cole (ed.), 315–332. New York: Academic Press.
- Vanderveken, Daniel. 1990. "Meaning and speech acts." *Principles of Language Use*. Cambridge: Cambridge University Press.
- doi Von Stechow, Ralf and Gillies, Anthony S. 2010. "Must... stay... strong!" *Natural Language Semantics* 18 (4): 351–383.
- doi Wathen, Nadine C. and Burkell, Jacquelyn. (2002). "Believe it or not: Factors influencing credibility on the web." *Journal of the American society for information science and technology* 53 (2): 134–144.
- Weber, Max. 1919. *Le savant et le politique*. Paris: Plon.
- doi Willett, Thomas. 1988. "A cross-linguistic survey of the grammaticization of evidentiality." *Studies in Language*.
- Williams, Bernard. 2010. *Truth and truthfulness: An essay in genealogy*. Princeton University Press.
- doi Williamson, Timothy. 1996. "Knowing and asserting." *The Philosophical Review* 105 (4), 489–523.
- Wolf, Lavi. 2015. Degrees of assertion. Ph. D. thesis, Ben-Gurion University of the Negev.
- Wright, Crispin. 1994. *Truth and objectivity*. Harvard University Press.

-  Yalcin, Seth. 2007. "Epistemic modals." *Mind* 116 (464): 983–1026.
- Yalcin, Seth. 2024. "Defining common ground." *Linguistics and Philosophy*, 47(6) : 1045–1070.
- Yatsushiro, Kazuko, Trinh, True, Z'ygis, Marthena, Solt, Stephanie, Benz, Anton and Krifka, Manfred. 2022. "Certainly but not certain: The expression of subjective and objective probability." *Glossa: a journal of general linguistics* 7 (1).
-  Zahra, Kiran, Imran, Muhammad and Ostermann, Frank. O. 2020. "Automatic identification of eyewitness messages on twitter during disasters." *Information Processing and Management*.