

#Soundwordsmatter: Epistemic modality and evidentiality in Twitter discourse on racism

Claudia Borghetti & Ana Pano Alamán
University of Bologna

Abstract This paper presents the results of a corpus-based pilot study aiming to investigate whether and to what extent epistemicity and evidentiality contribute to the degree of assertiveness of discourse on Twitter. The analysis was conducted on 900 tweets in English, Italian, and Spanish, published between May and June 2020 after the killing of George Floyd in the USA. We explore the epistemic constructions employed by users to evaluate the object of discourse (racism, discrimination), and their use of evidentials for marking the source of information of their statements. The results show that epistemic markers are not frequent in the corpora, but that in most cases they display high commitment towards the truth of the users' propositions. As for evidentiality, its presence is even lower, especially in the English dataset. These findings suggest that controversial debates on Twitter favour the adoption of assertive and inferential strategies.

Keywords epistemic modality; evidentiality; racism; public opinion; Twitter.

1. Introduction

Many features of the discourse on Twitter have been investigated and identified in recent years. As an increasing number of linguistic and pragmatic multilingual studies have shown, the streams of messages published in the microblog are often managed by the users as conversational practices (Honeycutt & Herring 2009; boyd et al. 2010; Zappavigna 2012). Thus, tweets are highly context-dependent and elliptical due in part to the 280-character limit, but also to the fast-paced interactions taking place. Indeed, as with other social media platforms, discourse on Twitter has proved to be informal and personal in style, despite mostly being directed to unknown and unpredictable audiences in a public space (Scott 2015). Moreover, utterances usually show high degrees of assertiveness and intensification, especially when topics lead to polarisation within politically or ethically oriented threads (Conover et al. 2011).

The study presented in this paper aimed to investigate whether and to what extent ‘epistemicity’ (the truth-value speakers attribute to their propositions) and ‘evidentiality’ (the encoding of sources of information for such propositions) contribute to the degree of assertiveness of discourse on Twitter. The analysis was conducted on 900 tweets in English, Italian, and Spanish automatically extracted using the hashtags #racism, #racismo, and #razzismo, which were published between May and June 2020, following the killing of George Floyd in the USA and the protests and rallies that spread around the world afterwards (§3.1). The main goal was to explore the epistemic constructions employed by users to evaluate the subject of the discourse (racism, discrimination), and their use of evidentials for marking the source of information of their statements. While a few investigations have addressed epistemic modality and epistemicity on Twitter (Zubiaga & Ji 2013; Mulder 2018; Berg et al. 2020), to the best of our knowledge, none have analysed both epistemicity and evidentiality in a multilingual corpus of tweets.

2. Two linguistic categories

The relationship between epistemic modality and evidentiality is highly debated. Most researchers (Palmer 1986; Plungian 2001) believe that the latter is a sub-category of the former, as speakers’ marking of the source of the information for their statement contributes to convey the degree of confidence in what they say. Others consider epistemicity as a form of evidentiality (Chafe 1986). Still others (de Haan 1999, 2001; Nuyts 2001; Cornillie 2009; Hart 2010) conceptualise these two linguistic categories as distinct, yet closely linked; it is only epistemic modality that expresses an evaluation of the truth or probability of a proposition, while evidential expressions – when combined with epistemic constructions – specify on what evidence such evaluation is made. This last position is the one assumed in this study, as better illustrated below.

2.1 Epistemic modality

In linguistics, epistemic modality or epistemicity is the semantic category which refers to the “evaluation of the chances that a certain hypothetical state of affairs under consideration (or some aspect of it) will occur, is occurring or has occurred in a possible world” (Nuyts 2001: 21). In other words, it expresses the truth-value or degree of probability that a speaker attributes to a propositional content. The result of this evaluation can be conceptualised along a continuum, which goes from absolute certainty or one hundred per cent probability that

what is said is true (strong epistemic modality) to absolute certainty that it is not true (weak epistemic modality); in between these two extremes, the level of confidence in the statement decreases, going from high degrees of probability to less-likely possibility (Cornillie 2009; Hart 2010).

Epistemic modality is grammaticalised differently and to different degrees in specific languages (Palmer 1986; De Haan 1997). To limit our discussion to the three considered here, English, Italian, and Spanish share many properties and present a few significant differences. In all of them, modal verbs represent one of the most grammaticalised means to convey epistemicity (regarding Italian, see Pietrandrea 2004). In English, strong epistemic modality is often expressed through modal verbs like *must* or *will*, while weak epistemic modality can be conveyed by *may* or *could*, with other modals that encode intermediate levels of certainty (*should*). Italian and Spanish present comparable examples; on the other hand, these two languages employ additional highly grammaticalised markers, such as the morphemes used to realise epistemic future (*saldrá todo bien, andrà tutto bene*). Apart from this difference, in the languages considered, epistemic modality is mainly linguistically encoded by lexical means including adverbials (*certainly; probabilmente; quizá*), verbs (*to imagine; credere; suponer*), adjectives (*a 'clear' mistake; il 'presunto' colpevole; una 'posible' consecuencia*) and the relative constructions (*it is clear that; si presume che; es posible que*) (Martínez Caro 2004; Pietrandrea 2004; Hart 2010; González Ruiz et al. 2016). Interestingly, the absence of an epistemic modal verb or other markers does not mean that epistemic modality is not present; quite the opposite, in fact, as most often total certainty to truth is zero-marked (Marín Arrese 2004; Hart 2011).

As mentioned, epistemicity is linked to evidentiality. The two categories are often expressed by the same elements (e.g., in Germanic and Romance languages). Secondly, the latter influences the former, as one can show more or less confidence in what is said (epistemic modality or support), depending on the type and reliability of the source of information (evidential support) (De Haan 1997, 2001; Nuyts 2001). This bond is evident in the term *commitment*, which is often used (Palmer 1986; Pietrandrea 2018) to indicate the speaker's degree of certainty towards the truth of the proposition, including the nature of the evidence for their evaluation. While recognising the connections between the two categories, like others (Rubin 2007; Cornillie 2009), we conceptualise epistemic commitment in a narrow sense, limiting its meaning to epistemicity only.

2.2 Evidentiality

Evidentiality is defined here as the linguistic category which expresses on what basis or source of information a person attributes a truth-value to a proposition. “That is, evidentials [or evidential markers] indicate how the speaker has come to know what they are claiming” (Hart 2011: 758). In this sense (see also Willett 1988), evidentials contribute to marking an epistemic attitude towards a state of affairs only ‘indirectly’. While they “indicate that there are reasons for the assumption made by the speaker” (Cornillie 2009: 57) and can serve as strategies to legitimise it (Hart 2010), evidential markers do not convey an epistemic evaluation per se.

In this narrow sense, various types of evidentials have been identified (Chafe 1986; Willett 1988; Bednarek 2006; Pietrandrea 2018). Generally, in traditional classifications, they are commonly conceptualised according to a macro-distinction between direct and indirect evidence. Evidential markers are direct when speakers base their statements on first-hand (visual, auditory, other sensory) sources of information; they are indirect when the speaker’s knowledge is based on inference (‘inferentials’) or what they have heard from others (‘reportive’ or ‘quotative’ evidentials) (Willett 1988). Like other West European languages, English, Italian, and Spanish mainly mark evidentiality lexically. In English, for instance, examples of evidential markers are: *I have read* *it is going to rain* (direct (visual) evidence), *it seems* or *apparently* (indirect inferential), *they say* or *according to* (indirect-reportive). While some of these are relatively unproblematic, many forms have multiple meanings, some of which are epistemic. This is especially true of inferentials (Auwera & Plungian 1998), as shown by the sentence *There are, it is said/it seems, many victims*, where the constructions *it is said/it seems* convey both epistemic (uncertainty) and evidential meanings (attribution of the information to hearsay) (Dendale 1993; Dendale & Tasmowski 2001: 345). As we will highlight, these conceptual issues impact considerably on any attempt to empirically investigate epistemicity and evidentiality.

With specific reference to the type of discriminatory discourse that we focus on in this study, it has been pointed out that evidentiality represents an important means “by which speakers, in order to overcome the epistemic safeguards of their audience, offer ‘guarantees’ for the truth of their assertions” (Hart 2011: 757-758). Various types of evidential markers can be used to put this legitimisation strategy in place. Expanding and revising Bednarek’s (2006) classification, Hart (2011) lists six strategies which offer bases of knowledge

and thus give legitimacy to propositions. He arranges these forms of evidence on a decreasing scale of reliability, from the most objective and therefore most reliable (1) to the most subjective and thus least reliable (6): (1) “Perception” – which mostly corresponds to direct evidence in Willett’s categorisation (1988) – provides directly attested sensory evidence (*it appears; visibly*); (2) “Proof” implies giving some sort of testimony based on attested results (*many studies show; statistics say*); (3) “Obviousness” constitutes indirect evidence inferred from reasoning or self-evidence (*obviously; it is manifest*); (4) “Public (or general) knowledge” marks the propositional content as based on what speakers consider part of the epistemic background shared with their audience (*it is well known; famously*); (5) “Expert knowledge” provides external support for the speaker’s claim by attributing it to assumed expert sources. This form of legitimation is realised through direct or indirect quotations introduced by specific markers (*X said; as stated by*); (6) “Epistemic commitment” – or rather ‘epistemic authority’ – bases evidence on the speaker’s self-attested qualified knowledge of the issue at hand (*being a member of the community; having a background in sociology*).

3. Public debate on Twitter

Launched in 2006, the microblog Twitter allows users to publish in their profiles messages of no more than 280 characters, which will be visible to other users in an aggregated timeline. Users may also retweet other users’ messages or reply to other tweets, raising the visibility of the shared information (boyd et al. 2010). They can also include in their tweets links to sources outside the microblog, one or more mentions (@username) to other users, and #hashtags, which permit the labelling or tagging of the topic of the message, and connecting it to messages that include the same hashtag, thus reaching wider audiences and fostering “ambient affiliation” (Zappavigna 2012).

As pointed out in numerous studies (Carney 2016; Indrawati 2021; Gaisbauer et al. 2021), social media, and Twitter in particular, are powerful communication platforms; they may be considered as a venue for open discussion of social and political issues, and as a space for social movements to take action. Indeed, since they allow differing points of view to reach such a vast public, social media have been considered as “counter-public spaces”, an expansion of Jürgen Habermas’ (1989 [1962]) concept of “public sphere”. Fraser defines them as “parallel discursive arenas where members of subordinated social groups invent and circulate counter discourses to formulate

oppositional interpretations of their identities, interests, and needs” (1992: 67) in response to hegemonic publics (Gaisbauer et al. 2021). As Carney (2016) explains, increased access to the Internet and the ability to read and contribute to discussions on social media via mobile phones allow people to integrate the public sphere into their daily activities. Thus, participation in online communities fosters different connections and enables users “to participate in campaigns and social movements, and to exchange opinions in social media in their own ways and language, drawing upon personal experiences, knowledge, engagements, values and judgements” (Rasmussen 2016: 79).

In this study, we are concerned with public debate on Twitter around racism, which, according to Smedley (1998), is a behaviour or belief representing the racial worldview that inherited physical characteristics and characteristics of identity are related. In the last decade, social media have provided new arenas for conversations about race and racial inequality particularly in the USA; in fact, two of the most used hashtags in Twitter’s history focus on race and criminal justice (Anderson 2016).¹ The debates around racism and police brutality against black individuals and communities in the USA that have taken place on Twitter since 2013 are often heated exchanges that “reveal the emergence of a few dominant ideological positions, emphasizing how different groups viewing the same media coverage interpret issues of race and police violence in dramatically different ways” (Carney 2016: 3; see also Smith et al. 2014). Since the platform does not allow space for a lengthy and nuanced conversation to unfold, a central characteristic of the debate is making an immediate impression on other users by adopting different strategies and making a purposeful use of the platform affordances, such as hashtags.

3.1 #BlackLivesMatter and the George Floyd case

Black Lives Matter (BLM) is a social movement condemning violence towards black people. The hashtag #BlackLivesMatter² represents the movement and has been used in social media as a call to action. Through information sharing

¹ According to Twitter, #Ferguson was the most used social-issue hashtag in the 10-year history of the microblog, while #BlackLivesMatter was third. The hashtag is considered to be a central strategy of the BLM movement.

² The hashtag was created in 2012, after the murder in Florida of Trayvon Martin, but it went viral in 2014 after the deaths of Michael Brown and Eric Garner. Although it has been recognized as an important hashtag of social change, it has also attracted criticism and resistance (Wilkins et al. 2019).

and a widespread use of the hashtag, BLM has managed to build an audience to promote its topics of interest, and to engage with others online (Bryan 2016). An analysis conducted by the Pew Research Center (Anderson 2016) on public tweets containing the hashtag (2013-2016) shows that the volume of race-related tweets tends to peak in the immediate aftermath of high-profile events, usually reflecting a synthesis of reactions rather than an account of the details of those events. This was particularly manifest after the killing of George Floyd in 2020. Floyd was an African-American man who was murdered by a white police officer, Derek Chauvin, in Minneapolis, Minnesota, during his arrest on May 25th. After this event, protests and rallies against police brutality towards black people spread quickly in the US and globally. At the same time, the hashtag #BlackLivesMatter gained momentum on Twitter, becoming a trending topic, that is, one of the most discussed topics in the microblog that week.

4. Method and data collection

Linguistic and pragmatic studies in different languages (Scott 2015; Spina 2019; Pano Alamán 2019, among others) have shown that tweets are often elliptical, dialogic and highly context-dependent, especially when the conversations around current political and social issues display a strong polarization following dramatic events such as the killing of George Floyd.

Given these assumptions, it was expected that our analysis would lead to the identification of a substantial number of linguistic choices showing a high degree of certainty as well as intensification in all three corpora. We intended to investigate the specific role played by epistemicity and evidentiality in shaping the assertive tone of argumentative discourse related to racism on Twitter. Accordingly, the following research questions were formulated:

1. RQ1. To what extent do epistemic constructions and evidentials contribute to making the users' utterances assertive in this context?
2. RQ2. What similarities and differences, if any, do the three corpora present in terms of speakers' epistemic commitment and evidential marking?

This is a corpus-based pilot study. We tested on the corpus the validity of the linguistic forms and structures derived from the theoretical issues discussed in Section 2 about epistemic modality and evidentiality. As we annotated the corpus, we noted some elements that we had not foreseen at the start. Therefore, a complementary corpus-driven approach was adopted for the analysis of tweets. Unlike a corpus-based investigation, where existing linguistic

patterns can be tested and validated, corpus-driven research is more inductive, so that the linguistic constructs emerge from a qualitative analysis of the corpus (Tognini-Bonelli 2001; see also Hunston 2011).

Regarding the corpus design, we restricted the data collection to one platform, Twitter, and decided to gather thematically organised streams of online discourse (hashtagged tweets) (Androutsopoulos 2013). Thus, the data were automatically extracted using the hashtags #racism, #racismo, and #razzismo, assuming that most of the debate on the microblog following the killing of George Floyd in Minneapolis would focus on this aspect. Secondly, we limited the collection to a specific period of time, since this approach “reconstructs a shared context for the tweets” (Kreis 2022: 83). The tweets were published in English, Spanish, and Italian from May 25th to June 1st, 2020, when the first protests and demonstrations against police brutality towards black people and racism took place in many sites around the world. In order to retrieve these tweets, which were extracted in March 2021, we used Vicinitas, a software that works with Twitter’s API. We used the *Historical Tweets* service provided by the software, which allowed us to search for messages related to trackers (hashtags) posted in the past, in this case, May-June 2020. These resulted in more than 10,746 tweets, which were downloaded in three Excel sheets, one for each language. As Table 1 shows, the number of tweets containing these key hashtags during the period considered was high, especially those in English, but also in the Spanish data. According to the metadata extracted with the tweets, most of these were posted from the USA and (for those in Spanish) by Hispanics residing in the USA, though other messages published in these two languages are geolocalized in other countries. The tweets posted in Italian, which are numerically inferior, all came from users located in Italy. Finally, in order to establish a proportion for a more reliable quantitative and qualitative comparison between datasets, we selected the first 300 tweets from each, obtaining a multilingual corpus of 900 tweets (26,426 tokens).

Hashtags	N. tweets	N. selected tweets
#racism (EN)	8,153	300 (9,529 tokens)
#racismo (ES)	2,202	300 (8,667tokens)
#razzismo (IT)	391	300 (8,230 tokens)
TOTAL	10,746	900 (26,426 tokens)

Table 1: Corpus composition

4.1 Data analysis

An annotation scheme based on previous studies (Blakemore 1994; Hart 2010; González Ruiz et al. 2016; Pietrandrea 2017, 2018) was first developed to code and to analyse the linguistic indicators of epistemic modality and evidentiality within the three languages.

For epistemic modality, for example, the analysis grid included adverbials (1) and morphemes like those in reportive conditional (2) and epistemic future (3) included in the examples below extracted from the dataset:

- (1) [...] *maybe* all of you need to understand what that means.³
- (2) *L'autopsia su #GeorgeFloyd escluderebbe segni di asfissia traumatica.*
- (3) *Così dovrà finire e finirà l'impero degli #USA, nel rogo del #razzismo e della #barbarie.*

Other indicators of epistemic modality considered in the multilingual annotation scheme were modal verbs such as “will” (4) and “poder” (5), where the latter is used in combination with an epistemic morpheme to realise “podría”:

- (4) *Eventually, doctors will find a coronavirus vaccine, but black people will continue to wait [...].*
- (5) *#GeorgeFloyd se ha convertido, lamentablemente, en un mártir, quien pone a luz pública que aún podría haber una batalla más seria que combatir en el mundo [...].*

One last example taken from the analysis grid concerns epistemic complement-taking predicates (6 and 7):

- (6) *#InHumanos claro que fue un acto de #Racismo [...].*
- (7) *Pensavamo che il Covid fosse il peggio che ci fosse capitato [...].*

Besides epistemic modality, the annotation scheme served to codify evidentials. Again, the following are examples taken from the corpora under investigation, which served as guidelines during the data coding. Some of the relevant categories were direct forms of evidence (“perception”, according to Hart

³ These sentences are taken from several tweets of the three sub-corpora; in this section, tweets are only partially reproduced following the regular citation rules.

2011) (8), reported or quotative (“expert knowledge”) (9 and 10) and inferred evidence (“obviousness”) (11):

- (8) [...] *Mai visto nessuno protestare per questo motivo.*
- (9) *Lo dijo el Rey Lebron y es la verdad.*
- (10) *From @nike: Don't pretend there's not a problem in #America [...].*
- (11) *El #racismo no es más que la soberbia de un ser evidentemente inferior [...].*

To manage the conceptual issue of distinguishing epistemic markers from evidentials (see §2), we decided to consider as evidentials only those forms whose legitimising value (i.e., conveying the source of information for a statement) clearly predominated over the epistemic value (Dendale & Tasmowski 2001). Another methodological decision taken was to annotate epistemicity only when an epistemic marker was used. This implies that we might have overlooked some instances of total certainty.

Once the annotation scheme and the related guidelines had been designed, all three corpora were analysed separately by the two authors. The process entailed several cycles of data coding and was accompanied by a progressive and collaborative refinement of the categories under investigation. Overall, these procedures were quite time-consuming, especially because the sources employed to delineate the preliminary version of the grid had not been originally intended to investigate epistemic modality and evidentiality on Twitter. In fact, Hart (2010) and González Ruiz et al. (2016) mainly address discourse in traditional media, and Pietrandrea (2017, 2018) dialogical spoken language.

In a third phase, we focused exclusively on the tweets which presented the phenomena investigated. We categorised these according to the epistemic support they presented, i.e. whether they conveyed +certainty or -certainty when validating the truth of the proposition expressed. We then classified all the data based on the presence or absence of evidential markers. Finally, a dedicated cycle of annotation was conducted on rhetorical questions, namely interrogatives which have the illocutionary force of assertions of the opposite polarity from what is seemingly asked (Ilie 1994; Han 2002), since they emerged as a stand-alone category of epistemic pragmatic markers across the three corpora, though with significant differences in terms of the number of occurrences.

5. A quantitative overview

This section presents a quantitative overview of the occurrences of the epistemic and evidential constructions identified within the three corpora. It helps answer the second research question of the study, namely “What similarities and differences, if any, do the three corpora present in terms of speakers' epistemic commitment and evidential marking?”.

As shown in Figure 1, strong epistemic modality prevails over weak epistemic modality in all three corpora, albeit to a slightly different extent. Therefore, no major differences were perceived between the English, Italian, and Spanish corpora. As anticipated (§4), a high degree of certainty was revealed. What is rather surprising is the overall configuration of results, namely that only a small minority of tweets present propositions whose (high or low) truth value is marked linguistically (on average, approximately 23 %).

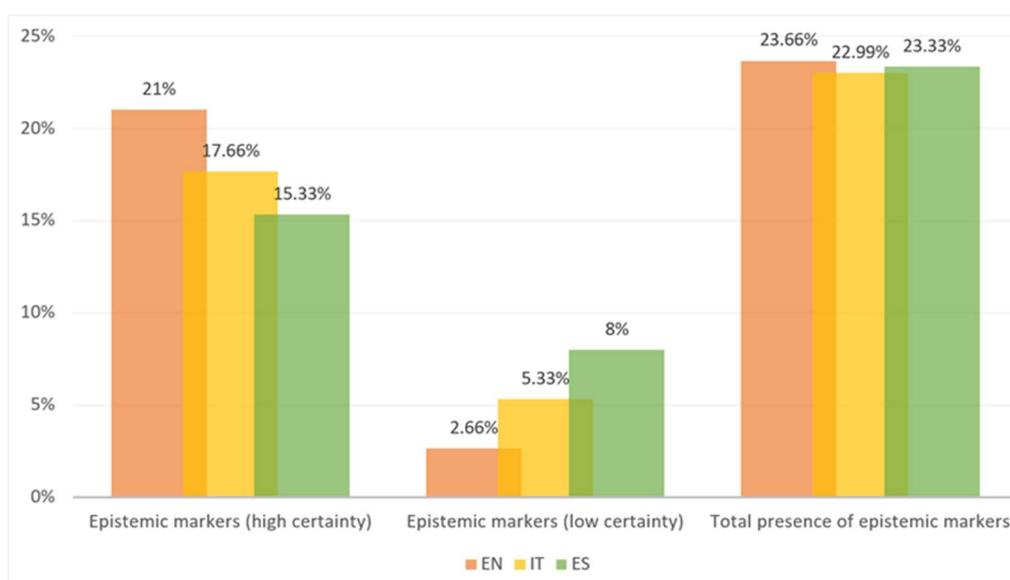


Figure 1: Percentage of tweets presenting epistemic markers in the three corpora

It is also interesting that, respectively 45.05% (EN), 20.26% (IT), and 22.84% (ES) of the total epistemic markers used in the tweets are rhetorical questions (Figure 2); however, the overwhelmingly greater use of this strategy should be noted in the English corpus.

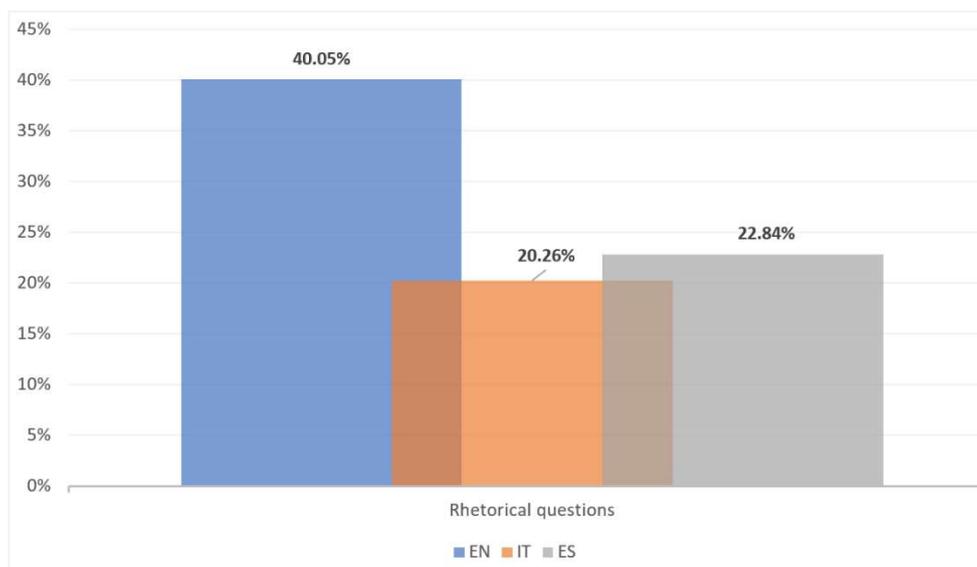


Figure 2: Percentage of rhetorical questions in the three corpora on the total of tweets (100%) presenting epistemic markers

Finally, as shown in Figure 3, just under half of the epistemic evaluations (46.49%) are accompanied by evidential justification in the English corpus; the percentage is higher in the Italian (62.63%) and Spanish (60%) corpora. In other words, when an indication of the degree of certainty is present (not often, as we have seen), the evidential source that justifies the validation of the truth may or may not be mentioned.

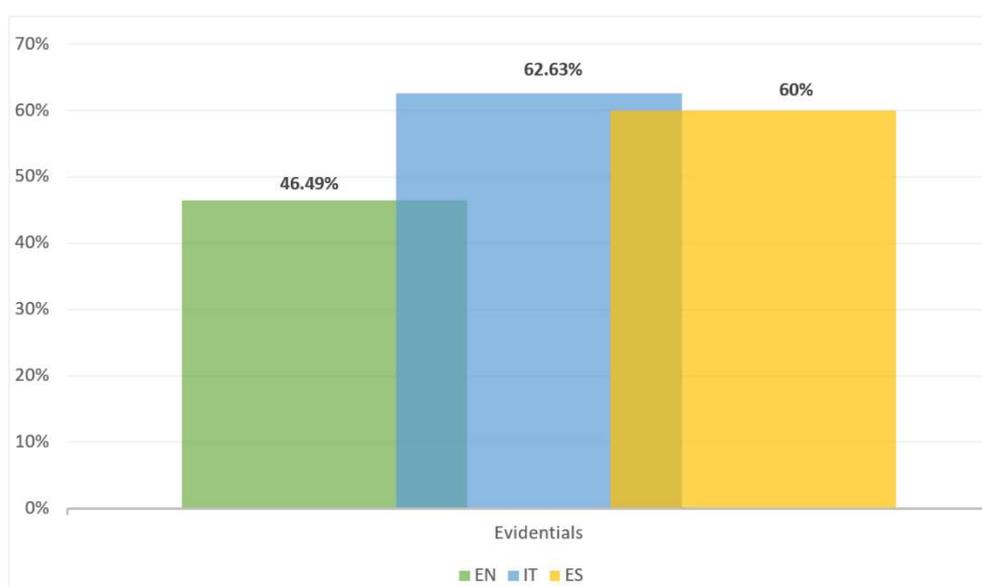


Figure 3: Percentage of evidentials in the three corpora on the total of tweets (100%) presenting epistemic markers

6. Epistemic modality: high vs. low degree of commitment

Tweets that contain at least one of the linguistic markers of epistemic modality with a value of high degree of commitment (Figure 1) are infrequent in the three datasets. As mentioned, *commitment* is intended here as the speaker's degree of certainty towards the truth of her/his proposition.

The results of the analysis show that users express certainty about what they say throughout epistemic adverbs, especially in the English dataset. See, for instance, the use of “clearly” in (12), commenting on the hashtag #BlackLivesMatter:

(12)



America's biggest issue is they had millions and millions of people vote for a racist piece of shit President and then the citizens post #BlackLivesMatter clearly they don't if you voted for him...give your head a fucking shake #PresidentTrump #RacistInChief #racism #america

Traducir Tweet

7:34 p. m. · 31 may. 2020 · Twitter for iPhone

(13)



Condemning #racism and #injustice should be everyone's business. Period. #BlackLiveMatter #condemninjustice #condemnracism

Traducir Tweet

8:10 p. m. · 31 may. 2020 · Twitter for Android

The author of the message states that the biggest issue in the US is that “millions and millions of people” voted for a “racist” President, Donald Trump. Assuming that many citizens using the hashtag may have voted for Trump, the author self-assuredly asserts that the meaning of the hashtag is therefore not valid. The epistemic marker invites their interlocutors to infer that ‘clearly, Black Lives do not Matter’ for those (implied ‘like you’) who voted for him. The adverb indicates “in a clear manner” (*Oxford English Dictionary*), so that the reason supporting the assertion appears to be straightforward, being based on their own ideology. Even though “clearly” may also have an evidential value, in this context it seems to signal the speaker's strong certainty more than a specific source of information.

In example (13), the modal structure “should + be” is used to assert that what everyone has to do is condemn racism. The noun “period” at the end of the message adds emphasis to this claim, implying that what the author says does not need to be discussed further. Depending on the context, epistemic uses of *should* can take on two contradictory values (Dufaye 2018). Here, it expresses high probability, as the author qualifies the condemnation of racism as normal and an obligation for all, according to their set of values.

We see similar patterns in the Spanish and the Italian datasets, where the use of what we coded as epistemic adverbs, epistemic phrases, and epistemic modal verbs, is also frequent. See the following tweets:

(14)



(15)



The adverbial phrase in Spanish “sin duda” (‘without doubt’) is used in (14) to claim emphatically that the killing of Floyd is clear sign of the world being sick. This linguistic element may be interpreted as reinforcing the truth value of the assertion in which it appears, since it presents a segment of information as evident and indisputable. But it also reflects a judgement that points to a extralinguistic reality “lo que ha ocurrido” (‘what happened’), revealing that the author is basing their statement on a dramatic perception of the events – see also the image embedded – and on a shared vision of the same events with their imagined audience on Twitter. Here, a combination of evidential strategies, based on both visual perception and public knowledge, is at work.

As for “Sicuramente sarà” in example (15), the author merges the adverb *sicuramente* (‘surely’) and the verb in the future tense *sarà* (‘will be’) with an epistemic value. They claim that Derek Chauvin, who is negatively described as a “lurido poliziotto” (‘filthy cop’), will not be incarcerated – instead, as the author of the embedded tweet expects to happen, he will be awarded a decoration and will be applauded for what he has done, implying that there will

be no justice for Floyd. As noted by Ranger (2011), *surely* is generally used to mark certainty, but, if combined with a verb in future tense, it may also mark disbelief or incomprehension. This might be explained as the consequence of the speaker's recognition of a wider discursive context. The message seems to implicitly refer to the non-indictment of other police officers, especially in the case of Eric Garner, killed by two policemen, who were not indicted after the decision of grand juries in 2014. As in this message, epistemic markers generally allow users to make clear statements about the plausibility of their conclusions, through marking the conclusion as such (Haßler 2010). Indeed, adverbs such as *clearly*, *sin duda*, or *sicuramente* characterise the implicit sources of the reported knowledge as given.

On the other hand, the results show a scarcer use of epistemic elements, such as *might be*, *apparently*, *mi pare che*, *quizá*, among others, which express a low degree of commitment towards what is said (see Figure 1). In these cases, the assertions have a different degree of likelihood to be true, as it can be seen in the following tweet:

(16)



Btw @amyklobuchar didn't reprimand any officers while in office. She need not to speak on behalf of Mr Floyd anymore & may be the reasons Chauvin was still on the force! #Racism must die Amy and it starts with holding officers accountable! No more trainings, we need convictions

Traducir Tweet

7:48 p. m. · 31 may. 2020 · Twitter for iPhone

In (16), “may be” introduces an argument that the reader of the tweet can accept as true or not. The sentence is connected to the first statement “@amyklobuchar didn't reprimand any officers”, so that the reader may infer that ‘probably’ the reason why the police officer who killed Floyd in Minneapolis was still on the force is that the Senator of Minnesota at that time, Amy Klobuchar, had not taken any measures against violent officers in the State.

We see identical values for the epistemic markers in Spanish and Italian, showing a low degree of certainty, as in these examples:

(17)



Quizá la próxima vez que un policía blanco decida apretar el gatillo, le venga la imagen de ciudades en llamas." excelsior.com.mx/global/miles-d...

10:00 p. m. · 1 jun. 2020 · Echobox

(18)



Vedo gente che #affonderebbe i #barconi e #insulta lo #straniero scrivere #ICantBreathe e post su #GeorgeFloyd. E a questo punto mi viene da pensare che lo fanno solo perché sono #fan di #SnoopDogg o #LeBronJames. Mi pare che abbiate le idee un po' confuse sul #razzismo

Traducir Tweet

9:01 p. m. · 27 may. 2020 · Twitter Web App

“Quizá” (‘perhaps’) conveys in (17) a reasoned suspension of the assertion according to which when another police officer decides to pull the trigger, he will see the image of a city on fire. Indeed, “quizá” reflects a doubt that arises in the mind of the speaker, who cannot provide any proof of what another ‘white policeman’, like Chauvin, will do in a similar situation. As for the syntactic construction in Italian “mi pare” (18), the completive sentence that follows the verb *parere* (‘to seem’): ‘that your ideas are a bit confused’, indicates extra-subjective evidence data or hearsay (‘I see people [writing on Twitter]’). The personal pronoun “mi” specifies that the speaker is the only one responsible for the judgement expressed in the utterance and for acquiring the knowledge (other users’ messages) on which the assertion is based.

In these tweets, the employment of low degree of commitment epistemic markers motivates the inference that they are willing to accept positions other than their own. However, if we look at the context where they are employed, we note that speakers often violate Grice’s maxim of quality, asserting something that they take for granted under a mechanism of uncertainty. This strategy may be connected to “hedging”, which refers to linguistic means used to indicate a lack of commitment to the truth value of a proposition, revealing disbelief, expressing caution, or even displaying an open attitude about a proposition (Rubin 2007). Therefore, the results allow us to consider these markers as signals of the presence of positions that cover a continuum from full commitment to confident commitment, based on hedging.

6.1 Rhetorical questions and irony

The analysis reveals a significant presence of rhetorical questions as linguistic elements that convey epistemic values in this context. As is well-known, these questions do not seek real answers but, on the contrary, provide an implicit answer. In the examples below, rhetorical questions conceal the statements that ‘we should watch no executions with our eyes closed’ (19) and that ‘beyond racism this may be a problem of inhumanity’ (20):

(19)



#Question ?
How many executions and murders should we watch with our eyes closed?
In the Whole World ????? #GeorgeFloyd 🙏 Rest in Peace
#NoWords #Revolution ... #NoChoice 🤔 #Pacifist is #Enough ... #Scandal #Horror #Racism #Stop
[facebook.com/1196393739/pos...](https://www.facebook.com/1196393739/pos...)
Traducir Tweet
8:16 p. m. · 31 may. 2020 · Twitter for iPhone

(20)



E se, aldilà del #razzismo, fosse un problema di disumanità?
Se penso a George Floyd, non penso ad un bianco che ha ucciso un nero.
Penso ad un uomo ucciso da un altro uomo. È lì, il nocciolo della questione.
#BlackLivesMatters #GeorgeFloyd
Traducir Tweet
11:54 a. m. · 28 may. 2020 · Twitter Web App

Ilie (1994) classifies rhetorical questions as mental-response-eliciting questions, claiming that they require a cognitive response linked to the interlocutor's acceptance of the answer implied by the speaker, so eliciting a mental recognition of its certainty or validity. But those identified in the three corpora also have the aim of strengthening assertions and making the tweet more memorable, especially through irony, as we see in (21), where the author rhetorically and ironically asks if the Ku Klux Klan is fashionable again.

(21)

Los casos de asesinato racial en EE.UU preocupan. Escudandose con "su ley" para matar personas de color es un abuso asqueroso! Y hay pruebas grabadas! El Ku Kux Klan estará de moda nuevamente?! 2020 - Más tecnología pero igual más retraso mental!
[#Racismo](#) [#JusticeForGeorgeFloyd](#)
 2:27 p. m. · 27 may. 2020 · Twitter Web App

6.2 Multimodal evidentiality

In their tweets, users make reference to news and articles, statements made by other users, and information appealing to shared knowledge. In the examples below, but also in (15), (17) and (19), users embed URLs linking to external data and to other tweets, and insert mentions to other users:

(22)

Truth. Plain and pointed. Read this [📄](#) [#RacialJustice](#)
[#race](#) [#COVID19](#) [#democracy](#) [#racism](#)
 Traducir Tweet
 · 31 may. 2020
 "African Americans have been living in a burning building for many years, choking on the smoke as the flames burn closer and closer. Racism in America is like dust in the air. It seems invisible — even if you're choking on it — until you let the sun in."
[latimes.com/opinion/story/...](#)
 8:01 p. m. · 31 may. 2020 · Twitter for iPhone

(23)

Secondo la polizia [#GeorgeFloyd](#) è morto CON un ginocchio in gola e non PER un ginocchio in gola. È il virus del razzismo.
 (wawe [@WaWe970](#))
[#28maggio](#) [#Minnesota](#) [#razzismo](#)
 Traducir Tweet
 8:30 p. m. · 28 may. 2020 · TweetDeck

Due to space constraints, we can make only a brief comment on these affordances, which allow us to explore epistemicity and evidentiality on Twitter within a broader, multimodal, context. Most messages in the datasets contain different types of what can be considered direct evidentials signalling sensory access (visual, textual) to discourse objects (Bergqvist & Kittilä 2020). More indirect evidentials express hearsay, as in (12) and (18), and assumptions based on a general shared knowledge of the world, as in (14), (15) and (16). In all

cases, they are used as a source of information, but, at the same time, they convey a judgement about their trustworthiness (Haßler 2010). To this end, users tend to gather data that generally confirm their own views and assumptions, while they ignore those with different views (Mancera Rueda & Pano Alamán 2020). Moreover, since Twitter algorithms offer personalised content and favour selective exposure, we may assume that most of these evidentials reinforce the users' own beliefs and world views on racism.

7. Discussion

Although the results obtained with this pilot study will need to be verified with further studies conducted on larger corpora, they help shed some light on the role played by epistemic and evidential constructions in shaping the assertive tone of public debates about racism on Twitter.

Our findings show that, contrary to what was expected (§4), the contribution of epistemicity to making the analysed tweets assertive is limited (RQ1). Even though markers manifesting high commitment towards the truth of propositions prevail over the contrary in all three corpora, less than a quarter of the messages present at least one epistemic marker, with minor differences among the languages under investigation. Two possible explanations are envisaged. The first is methodological in nature, as we codified epistemicity only when it was explicitly expressed by linguistic markers. It is however true that “total commitment to truth is zero-marked in most languages” (Marín Arrese 2004: 156). In fact, “total commitment of the speaker in non-hedged modality may be taken as evidence for the truth of their assertion on the assumption that the speaker is confident enough to make a categorical claim when they wouldn't want later to be undermined and lose credibility” (Hart 2011: 759). Thus, while the methodological decision we made allowed us to operationalise the theoretical construct of epistemic modality in a coherent manner for analysis purposes, it may have been the very reason why the number of Tweets presenting the speakers' high degree of commitment are fewer than expected in all three corpora.

The second explanation – which is linked to the first – is also theoretical in principle. Given the purposes of the study, we did not focus – either conceptually or analytically – on deontic modality and its relations with epistemicity. Yet this semantic category indicates “the degree of moral desirability of the state of affairs expressed in the utterance” (Nuyts 2006: 4), where morality is conceived widely to encompass what is legally permissible and

socially acceptable, according to the speaker's personal ethical criteria (Nuyts 2006; Charlow & Chrisman 2016). Thus, although investigating deontic modality would clearly have gone beyond the scope of this study, it might have helped explain why expressions like the following (24) sound assertive despite having no epistemic value.

- (24) [...] *we must be angry every time there is oppression of human beings, and we're very angry about #Racism*

Even though limited in terms of number, the epistemic constructions identified show some peculiarities which are worth noting. First, in all three corpora, they are signalled by recurring markers (RQ2). While some of these (e.g., “the real”, “clear”, “it's clear”, “certo”, “realmente”, “vero”, “es evidente”, “claro”, etc.) are attested in other types of texts (González Ruiz et al. 2016; Pietrandrea 2017, 2018), others seem to be linked to the very features of the conversations on Twitter. This is the case of pragmatic markers like interjections (“mah”, “bah”, “ya”, “period”) which often indicate epistemic constructions whose object of evaluation is syntactically independent (e.g., “è innocente. Bah!”). Second, as in this last example, they are in some cases ironical. When present, markers of uncertainty (“maybe”, “perhaps”, “it seems”, etc.) are employed with an ironic value (e.g., “It seems to me that you are a bit confused about #racism”); thus, they in fact convey +certainty rather than the opposite. Finally, particularly in the English corpus, we identified a substantial number of rhetorical questions, which have an epistemic status by definition, since they make claims or assertions, even if of the opposite polarity to that expressed by the question (Koshik 2005). In other words, these constructions are intended to induce the addressee to converge with the speaker's conclusions, and it is precisely this perlocutionary effect that contributes to attributing a truth value to the propositional content. In this regard, it is hard to explain why the English corpus has twice as many rhetorical questions compared to the other two corpora. We may hypothesise that, in the USA, public debates about racist episodes and the related popular and political/institutional responses have been taking place for a longer period of time; this may make US-based Twitter users more confident in conveying their high degree of commitment by means of linguistic choices which mostly rely on the audience's acceptance and validation of the truth-value of their propositions. Another possible explanation could be sought in the uneven distribution of rhetorical questions in public discourse across the languages considered. In newspaper commentaries, for example,

Dafouz-Milne (2008) in her study on the role played by metadiscourse markers in the construction and attainment of persuasion, identified twice as many rhetorical questions in the British *The Times* than in the Spanish *El País*.

As regards evidentiality, slightly more than half of the tweets (approximately 56% on average, across the three corpora) make explicit on what basis or source of information a truth-value is attributed to what the speakers say. It is arguable that a lack of evidentials contributes to making utterances assertive. However, in our data, this would be true only for a minority – albeit a substantial minority – of tweets. Again, a numerical difference is identified between the English corpus and the other two, as the speakers' assertions in the former are less often accompanied by evidential justification; legitimation strategies are employed in 46.49% of the English messages presenting epistemic markers but slightly more often (approximately 60% on average) in the Italian and Spanish ones. Again, this figure suggests that (USA-based) English speakers might feel less need to provide reasons as to why their audience should accept their assertions as true. In any case, across the three corpora, when bases of knowledge or sources of information are reported or implied, they are mostly introduced through either direct sensory evidence or indirect quotative strategies. The latter are quite varied in nature, ranging from “expert knowledge” (including references to pieces of news and public/official declarations) to “public knowledge” or hearsay (e.g., sayings and aphorisms). As described, most often these strategies assume the form of what we may call ‘multimodal evidentiality’, as the speakers provide evidential justification for their evaluations through the sharing of images, videos, and external links, as well as inserting mentions and hashtags in their messages.

8. Conclusion

This pilot study on Twitter conversations about racism has allowed us to test in this context the traditional classifications of epistemic markers and evidentials, which mostly concern written journalistic and academic texts, or oral conversations. The results have suggested that the existing taxonomies are valid for the detection of the linguistic elements which have epistemic and evidential functions in the tweets. However, some other features need to be integrated, as on Twitter most evidentials point to extra-linguistic information, and assertions with a high degree of commitment are highly implicit. Thus, for the future, we will develop a new grid of analysis and test it, until the interrater agreement is satisfactory.

Subjective, polemical, and emotional tweets about racism, which accumulate within fast-paced conversations on the platform, are elliptical and extremely fragmented since they contain different modes (images, videos) and affordances (mentions, hashtags, URLs). The technological and socio-situational specificities of the microblog, therefore, seem to affect the users' epistemic and evidential strategies. On the other hand, considering the percentage differences between the English corpus and the Spanish and Italian ones in the employment of highly inferential strategies, and the possible explanations provided in the discussion, it seems possible that the sociocultural background and context of users play a fundamental role too.

In this regard, future studies conducted on larger corpora are necessary to better understand what semantic and pragmatic values epistemic markers acquire in this context and what kind of verbal and non-verbal evidentials are employed by Twitter users to validate their opinions on controversial issues such as racism. It would indeed be very interesting to see in the future a study analysing, by means of more appropriate inferential statistical methodologies, how individual indicators of epistemicity and evidentiality, the different languages considered and the possible interactions between all these variables contribute to determining the greater or lesser degree of certainty of tweets on racism.

References

- Anderson, Monica. 2016. "Social Media Conversations About Race." *Pew Research Center*. 15 August 2016, <http://www.pewinternet.org/2016/08/15/social-media-conversations-about-race/> [last access on 30/09/2022].
- Androutsopoulos, Jannis. 2013. "Online data collection." In Christine Mallinson, Becky Childs, & Gerard Van Herk (eds.), *Data Collection in Sociolinguistics. Methods and Applications*, 236–249. London: Routledge.
- Auwers, Johan van der & Vladimir A. Plungian. 1998. "Modality's semantic map." *Linguistic Typology* 2, 79–124.
- Bednarek, Monika. 2006. "Epistemological positioning and evidentiality in English news discourse: A text-driven approach." *Text & Talk* 26(6), 635–660.
- Berg, Sebastian, Tim König, & Ann-Kathrin Koster. 2020. "Political opinion formation as epistemic practice: The hashtag assemblage of #metwo." *Media and Communication* 8(4), 84–95.

Bergqvist, Henrik & Seppo Kittilä. 2020. “Epistemic perspectives: Evidentiality, egophoricity, and engagement.” In Bergqvist, Henrik & Seppo Kittilä (eds.), *Evidentiality, Egophoricity, and Engagement*, 1–21. Berlin: Language Science Press.

Blakemore, Diane. 1994. “Evidence and modality.” In Ron E. Asher (ed.), *The Encyclopedia of Language and Linguistics*, 1183–1186. Oxford: Pergamon Press.

boyd, danah, Scot Golder, & Gilad Lotan. 2010. “Tweet, tweet, retweet: Conversational aspects of retweeting on Twitter”. *43rd Hawaii International Conference on System Sciences*. IEEE, <https://ieeexplore.ieee.org/document/5428313> [last access on 30/09/2022].

Bryan, Tacicia. 2016. *Black Lives Matter Toronto: A Qualitative Study of Twitter’s Localized Social Discourse on Systemic Racism*. Toronto. Ryerson University (MA Dissertation).

Carney, Nikita. 2016. “All lives matter, but so does race: Black lives matter and the evolving role of social media.” *Humanity & Society* 40(2), 180–199.

Chafe, Wallace L. 1986. “Evidentiality in English conversation and academic writing.” In Wallace L. Chafe & Johanna Nichols (eds.), *Evidentiality: The Linguistic Coding of Epistemology*, 261–272. Norwood: Ablex.

Charlow, Nate & Matthew Chrisman. 2016. “Introduction.” In Nate Charlow & Matthew Chrisman (eds.), *Deontic Modality*, 1–10. Oxford: Oxford University Press.

Conover, Michael, Jacob Ratkiewicz, Matthew Francisco, Bruno Goncalves, Filippo Menczer, & Alessandro Flammini. 2021. “Political polarization on Twitter.” *Proceedings of the International AAAI Conference on Web and Social Media*, 5(1), 89–96, <https://ojs.aaai.org/index.php/ICWSM/article/view/14126> [last access on 30/09/2022].

Cornillie, Bert. 2009. “Evidentiality and epistemic modality: On the close relationship between two different categories.” *Functions of Language* 16(1), 44–62.

Dafouz-Milne, Emma. 2008. “The pragmatic role of textual and interpersonal metadiscourse markers in the construction and attainment of persuasion: A cross-linguistic study of newspaper discourse.” *Journal of Pragmatics* 40, 95–113.

De Haan, Ferdinand. 1997. *The Interaction of Modality and Negation: A Typological Study*. New York: Garland.

De Haan, Ferdinand. 1999. “Evidentiality and epistemic modality: setting boundaries.” *Southwest Journal of Linguistics* 18, 83–101.

- De Haan, Ferdinand. 2001. "The relation between modality and evidentiality." *Linguistische Berichte* 9, 201–216.
- Dendale, Patrick. 1993. "Le conditionnel de l'information incertaine: Marqueur modal ou marqueur evidenciel?" In Gerold Hilty (ed.), *Actes du XXe Congrès international de Linguistique et Philologie romanes*, tome 1, 165–176. Tübingen: Francke.
- Dendale, Patrick & Liliane Tasmowski. 2001. "Introduction: Evidentiality and related notions." *Journal of Pragmatics* 33(3), 339–348.
- Dufaye, Lionel. 2018. "SHOULD in conditional clauses: When epistemicity meets appreciative modality." In Zlatka Guentchéva (ed.), *Epistemic Modalities and Evidentiality in Cross-Linguistic Perspective*, 52–66. Berlin: De Gruyter Mouton.
- Fraser, Nancy. 1992. "Rethinking the public sphere: A contribution to the critique of actually existing democracy." In Craig Calhoun (ed.), *Habermas and the Public Sphere*, 109–142. Cambridge (MA): MIT Press.
- Gaisbauer, Felix, Armin Pournaki, Sven Banisch, & Eckehard Olbrich. 2021. "Ideological differences in engagement in public debate on Twitter." *PLoS ONE* 16(3), 1–18.
- González Ruiz, Ramón, Dámaso Izquierdo Alegría, & Óscar Loureda Lamas (eds.). 2016. *La evidencialidad en español: teoría y descripción*. Madrid & Frankfurt: Vervuert/Iberoamericana.
- Habermas, Jürgen. 1989 (1962). *The Structural Transformation of the Public Sphere*. Cambridge: Polity.
- Han, Chung-hye. 2002. "Interpreting interrogatives as rhetorical questions." *Lingua* 112(3), 201–229.
- Hart, Christopher. 2010. *Critical Discourse Analysis and Cognitive Science: New Perspectives on Immigration Discourse*. London: MacMillan.
- Hart, Christopher. 2011. "Legitimizing assertions and the logico-rhetorical module: Evidence and epistemic vigilance in media discourse on immigration." *Discourse Studies* 13(6), 751–814.
- Haßler, Gerda. 2010. "Epistemic modality and evidentiality and their determination on a deictic basis: the case of Romance languages." In Gabriele Diewald & Elena Smirnova (eds.), *Linguistic Realization of Evidentiality in European Languages*, 223–248. Berlin: Mouton De Gruyter.

Honeycutt, Courtenay & Susan C. Herring. 2009. "Beyond microblogging: Conversation and collaboration in Twitter." *42nd Hawaii International Conference on System Sciences*, IEEE, <https://ieeexplore.ieee.org/document/4755499> [last access on 30/09/2022].

Hunston, Susan. 2011. *Corpus Approaches to Evaluation: Phraseology and Evaluative Language*. London: Routledge.

Ilie, Cornelia. 1994. *What Else Can I Tell You? A Pragmatic Study of English Rhetorical Questions as Discursive and Argumentative Acts*. Stockholm: Almqvist & Wiksell.

Indrawati, Desi. 2021. "Critical discourse analysis on representation of racism and solidarity in Adidas's tweets." *Lingua Cultura* 15(1), 109–119.

Koshik, Irene. 2005. *Beyond Rhetorical Questions: Assertive Questions in Everyday Interaction*. Amsterdam: John Benjamins Publishing.

Kreis, Ramona. 2022. "Data collection, preparation, and management." In Camilla Vásquez (ed.), *Research Methods for Digital Discourse Analysis*, 73–90. London: Bloomsbury.

Mancera Rueda, Ana & Ana Pano Alamán. 2020. *La opinión pública en la red. Análisis pragmático de la voz de los ciudadanos*. Madrid & Frankfurt: Iberoamericana Vervuert.

Marín-Arrese, Juana I. 2004. "Evidential and epistemic qualifications in the discourse of fact and opinion: A comparable corpus study." In Juana I. Marín-Arrese (ed.), *Perspectives on Evidentiality and Modality*, 153–184. Madrid: Editorial Complutense.

Martínez Caro, Elena. 2004. "Evidentiality and the verbal expression of belief and hearsay." In Juana I. Marín Arrese (ed.), *Perspectives on Evidentiality and Modality*, 185–204. Madrid: Editorial Complutense.

Mulder, Gijs, 2018. "(Yo) creo que as a marker of evidentiality and epistemic modality: Evidence from Twitter." In Ad Foolen, Helen de Hoop, & Gijs Mulder (eds.), *Evidence for Evidentiality*, 99–120. Amsterdam: John Benjamins Publishing.

Nuyts, Jan. 2001. *Epistemic Modality, Language, and Conceptualization: A cognitive-pragmatic perspective*. Amsterdam: John Benjamins Publishing.

Nuyts, Jan. 2006. "Modality: Overview and linguistic issues." In William Frawley (ed.), *The Expression of Modality*, 1–26. Berlin: Mouton De Gruyter.

Oxford English Dictionary. [Online]. 2022. “Clearly”.
<https://www.oed.com/view/Entry/34093?redirectedFrom=clearly&> [last access on 30/09/2022].

Palmer, Frank R. 1986. *Mood and Modality*. Cambridge: Cambridge University Press.

Pano Alamán, Ana. 2019. “Condensación y fragmentación del discurso político en Twitter”. In Nel·lo Pellisser Rossell & Joan Oleaque (eds.), *Mutaciones discursivas en el siglo XXI: la política en los medios y las redes*, 75–92. Valencia: Tirant lo Blanch.

Pietrandrea, Paola. 2004. “L’articolazione semantica del dominio epistemico dell’italiano.” *Lingue e linguaggio* 2, 171–206.

Pietrandrea, Paola. 2017. “Le costruzioni epistemiche dialogiche dell’italiano. Una modellizzazione corpus-driven.” In Cristiana De Santis & Nicola Grandi (eds.), *CLUB Working Papers*, 90–117. Bologna: AMS Acta Università di Bologna.

Pietrandrea, Paola. 2018. “Epistemic constructions at work. A corpus study on spoken Italian dialogues.” *Journal of Pragmatics* 128, 171–191.

Plungian, Vladimir A. 2001. “The place of evidentiality within the universal grammatical space.” *Journal of Pragmatics* 33, 349–357.

Ranger, Graham. 2011. “Surely not! Between certainty and disbelief.” *Discours* 8, 3–22.

Rasmussen, Terje. 2016. *The Internet Soapbox. Perspectives on a Changing Public Sphere*. Oslo: Universitetsforlaget.

Rubin, Victoria L. 2007. “Stating with certainty or stating with doubt: Intercoder reliability results for manual annotation of epistemically modalized statements.” *Proceedings of NAACL HLT 2007*, 141–144. Rochester (NY): Association for Computational Linguistics.

Scott, Kate. 2015. “The pragmatics of hashtags: Inference and conversational style on Twitter.” *Journal of Pragmatics* 81, 8–20.

Smedley, Audrey. 1998. “Race and the construction of human identity.” *American Anthropologist* 100(3), 690–702.

Smith, Marc A., Lee Rainie, Itai Himelboim, & Ben Shneiderman. 2014. *Mapping Twitter Topic Networks: From Polarized Crowds to Community Clusters*. Washington, DC: Pew Research Center.

Spina, Stefania. 2019. *Fiumi di parole. Discorso e grammatica delle conversazioni scritte in Twitter*. Roma: Aracne.

Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work*. Amsterdam: John Benjamins Publishing.

Willett, Thomas. 1988. "A crosslinguistic survey of grammaticalisation of evidentiality." *Studies in Language* 12, 51–97.

Wilkins, Denise, Andrew G. Livingstone, & Mark Levine. 2019. "Whose tweets? The rhetorical functions of social media use in developing the Black Lives Matter movement." *British Journal of Social Psychology* 58(4), 786–805.

Zappavigna, Michele. 2012. *Discourse of Twitter and Social Media: How We Use Language to Create Affiliation on the Web*. Bloomsbury: London.

Zubiaga, Arkaitz & Ji Heng. 2013. "Tweet, but verify: Epistemic study of information verification on Twitter." *Social Network Analysis and Mining* 4(1), 1–12.