



Contents lists available at ScienceDirect

Journal of Pragmatics

journal homepage: www.elsevier.com/locate/pragma

Using prosody to express evidentiality. The case of the quotative

Fernando Bermúdez

Department of Modern Languages, Uppsala University, Box 636, 751 26, Uppsala, Sweden



ARTICLE INFO

Article history:

Received 19 October 2022

Received in revised form 18 June 2023

Accepted 20 June 2023

Available online 22 July 2023

Keywords:

Evidentiality

Prosody

Quotative

ABSTRACT

The goal of this study was to investigate if pure prosodic marking of evidentiality, more specifically quotation, does exist and, secondly, if the particular prosodic traits involved in such prosodic marking are arbitrary or somewhat motivated. Data containing a set of identical sentence pairs in Spanish (the only difference being the presence or absence of quotation in them) was collected by means of a Discourse Completion Task and analyzed in order to find prosodic patterns. The results suggested that an affirmative answer must be given to the first question, as the quoted fragments were consistently marked in the four languages with a break before the quoted fragment (and often after it) and a change in voice quality, more specifically by raising the pitch of the snippet. An extension of the study was carried out in three different languages from three different language families (Swedish, Polish and Hungarian) to check if the traits observed were an idiosyncratic feature of Spanish or a more widespread phenomenon. The same conclusive results were observed in all three languages. Regarding the question of the arbitrariness or motivation of the prosodic markers, the results suggest that both of the prosodic traits found are indeed motivated. Speakers use metonymically traits that are the most salient when in an everyday conversation a new speaker takes the floor (a gap and a change of voice quality) to convey that the enunciator has changed and the fragment at glance is to be attributed to an individual other than the speaker. The consistent choice of pitch raising to mark the change of enunciator (as opposed to pitch lowering, for example) is argued to be related to issues of language embodiment.

© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Evidentiality (the linguistic coding of the source of information speakers have for their statements) is a field of research that has been gaining popularity in the last 30 years. Nevertheless, almost without exception, studies on evidentiality have been limited to the study of morphological and lexical markers (Chafe and Nichols, 1986; Aikhenvald, 2004; Aikhenvald and Dixon, 2014; González Ruiz et al., 2016, among many others), without any reference to their prosody. This is true also for studies that take an interactional approach (Squartini, 2012; Nuckolls and Michael, 2014; Söderqvist, 2020). It is not until the very last years that some isolated works have appeared that highlight the possible role of prosody in the expression of evidentiality (Estellés Arguedas and Albelda Marco, 2014; Estellés Arguedas, 2015; Vanrell et al., 2017; Cabedo Nebot and Cornillie, 2018). Such studies, though, have been very specific, conducted from different theoretical perspectives, and without an explicit goal of getting insights into the relation between evidentiality and prosody.

In an attempt to fill this gap, and shed new light on both evidentiality and prosody, the present article aims to investigate the role of prosody in the expression of evidential meanings, based on the analysis of spoken Spanish data.

E-mail address: fernando.bermudez@moderna.uu.se.

<https://doi.org/10.1016/j.pragma.2023.06.009>

0378-2166/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

The research questions that we will attempt to answer here are:

- 1) Are there “prosodic evidential markers”, in the same way that there are evidential particles, adverbs or suffixes, in other words, prosodic patterns that by themselves mark a statement with an evidential value?
- 2) If there are “prosodic evidential markers”, are their prosodic features arbitrary or are they somehow related to the meaning the markers convey?

In this article, a particular evidential marking is investigated, namely the quotative.

2. Evidentiality and related notions

The term *evidentiality* itself is relatively new - if for example we compare it with the millenary history of modality. In fact, it appears for the first time in a work by Boas (1947). Moreover, the term and the notions linked to it did not initially have a direct impact on linguistic research (with exceptions, such as Jakobson, 1957), but were rather seen primarily as a specific grammatical category of certain languages, such as the languages of the Americas. Partly due to the publication of Chafe and Nichols (1986), evidentiality has gradually become a research field in its own right within linguistics, and gradually it has been recognized that all languages have the possibility of expressing the source of information, using more or less grammaticalized means (Bermúdez, 2004, 2006; Squartini, 2008; Guentchéva, 2018).

During the years since the publication of Chafe and Nichols (1986), various controversies have arisen around the concept of evidentiality (see Dendale and Tasmowski, 2001 and Aikhenvald, 2004 for a review). The three most important are the relationship between evidentiality and epistemic modality, the scope of the notion of evidentiality itself, and its internal structure.

Many scholars have claimed that evidentials are means for expressing the speaker's epistemic attitude. Palmer (1986), for example, claims that speakers encode their degree of confidence or commitment regarding what is said both through expressions that directly express certainty or doubt (modals) and through expressions that encode the way in which the speaker gained access to such information (evidentials). Similar views have been defended by Chafe (1986) and Frawley (1992), among others. More recently, however, the prevalent opinion is that modality and evidentiality are two independent categories, though related in many ways (Lazard, 1999; Aikhenvald, 2004; de Haan, 2005; Guentchéva 2018). In this article, we adopt the view that evidentiality is distinct from the expression of the epistemic attitude of the speaker, i.e., from the speaker's degree of certainty or commitment regarding what is said, even if speakers can use evidential markers (and many other devices, like demonstratives, to name just one) to convey modal meaning.

As for the scope of the notion of evidentiality, there has been debate whether to include notions such as *mirativity* (the notion of “unprepared mind” or “sudden realization”), *egophoricity* (speaker involvement or epistemic authority) and *engagement* (the distribution of knowledge between the participants in an interaction) in the domain of evidentiality (see de Lancey, 2001; Lau and Rooryck, 2017 and Bergqvist and Kittilä, 2020 for a detailed discussion).

Regarding the internal structure of evidentiality, some semasiological proposals have been made, largely based on distinctions made by so-called evidential languages, i.e. those languages with very specialized morphological systems for the expression of evidential meanings.¹ The best known is that of Willett (1988: 57, see Fig. 1 below), which identifies the distinction between direct and indirect evidence as the primary one (see Aikhenvald, 2004; Plungian 2010, and Izquierdo Alegría, 2016 for alternative classifications):



Fig. 1. The category of evidentiality according to Willett (1988).

¹ Some scholars, such as Aikhenvald (2004), prefer to restrict the use of the term *evidentiality* to the grammatical category found in those languages, and to use ‘evidentiality strategies’ for all other means to express evidential meanings.

Models like Willett's work acceptably well to describe the highly grammaticalized systems of evidential languages, like Tariana or Tuyuca. They have, however, trouble describing the expression of evidential meanings in languages such as English or Spanish, and have certain conceptual and empirical limitations when applied as a description of the semantic domain of evidentiality. The evidential value of quotative, the one that this study focuses on, is not included in Willett's classification, but it could find its place along with second-hand evidence in his model.

As for the possible relationship between evidentiality and prosody, there is still much to explore, though some attempts have been made in the studies already mentioned. Cabedo Nebot and Cornillie (2018), for example, have found some connections between intonation and (inter)subjectivity in certain evidential adverbs, while Vanrell, Armstrong & Prieto (2017) have found evidential undertones in specific variations on the intonation in polar interrogatives in Catalan. Additionally, García Negroni & Caldiz (2016), in a pioneer work done within a Ducrot-inspired framework, suggested that intonation plays a significant role in the codification and decodification of evidentiality, since the prosodic marking "may itself perform that modal function". They work, though, with a broader definition of evidentiality, as the "enactment of viewpoints that establish an interactive connection with discourse frames rooted on previous or potential discourses". Further pursuing the role of intonation in the expression of so called evidential viewpoints, Caldiz (2018), within the same theoretical frame, emphasizes the role of circumflex in expressing the point of view of the speaker in River Plate-Spanish, arguing that the circumflex tone establishes a dialogical bond with discursive frames on which utterances are rooted, thus unveiling self's argumentative stance and fostering other's interpretation, while Estellés Arguedas (2015) identify certain "special prosodic configurations" in direct reported speech situations, and that they are more overt when the utterance lacks a *verbum dicendi*. A difference between our work and these works within the Ducrotian framework, is that they tend to equal reported speech with evidentiality, which we do not, as we argue for in the discussion of examples (3)–(6) below. However, no comprehensive work has been published on the relationship between evidentiality and prosody.

The prosody of reported speech, however, has received greater attention in the literature, even though those studies do not link their results to a broader discussion on prosodic marking of evidentiality.² Klewitz and Couper-Kuhlen (1999), for example, found prosodic marking of reported speech like global pitch and loudness shifts and rhythm changes in English. Nevertheless, they claim that those traits constitute a stylistic device rather than a norm: according to them, they may be used or not, depending on speakers' local goals and strategic choices. Günthner (1999), from a point of view based on Bakhtinian polyphony, discusses the ways in which speakers express their evaluative perspective on reported utterances through prosodic clues, such as pitch movements, volume variations and lengthened vocals. More recently, Kasimir (2008) conducted an experiment where speakers were asked to read aloud written sentences containing sub-clausal quotation marks and subsequently another set of informants were instructed to listen to the recorded sentences and typewrite them down, to see if they use quotation marks to signal reported speech. She did not conduct a phonetic analysis of the recordings, which makes a comparison with our results difficult; nevertheless, she did not find any compelling correlation between the presence or absence of quotation marks in the original written examples and the presence or absence of quotation marks in corresponding back-translations from oral renditions, which suggests that the prosodic clues are not fully conventionalized devices. Bolden (2004), from a conversational analysis framework, presents a comprehensive picture of how reported speech is carried out in Russian. However, she does not focus on the prosodic marking of reported speech itself but rather on how the marking (mostly lexical or grammatical but even prosodic) extends beyond the limits of a turn conversational unit and how the end of a quote is marked. In those cases where she finds prosodic traits aligned with other framing devices separating reported speech from other discourse, those prosodic traits are: pitch changes (both pitch rising and lowering), pauses, volume changes, and speed changes, which can be compared with our results. However, a common characteristic of these studies is the conclusion that the relationship between prosody and reported speech neither is a norm nor fully conventionalized.

3. Evidentiality and deixis

In this article, we assume a view of evidentiality as a deictic category in the sense that it locates the (source of) information in relation to the speaker and the speech situation, as stated in Bermúdez (2006, 2011, 2016; see Frawley, 1992 and Mushin, 2001 for previous models based on deixis).

Spatial deixis, that is, the location of objects in space in relation to a deictic centre, can be defined by three continuous and independent parameters: *direction*, *distance*, and *reference points*, one of which, the deictic centre, usually the speaker, is of vital importance, since it is the reference point that creates the perspective. Thus, an expression like "behind x" can be described as "in the *same direction* as the *reference point* x (from the point of view of the deictic centre), but at a *greater distance*" (see Fig. 2).

² In Section 4 below I will discuss briefly the differences between reported speech and reportative evidentiality.

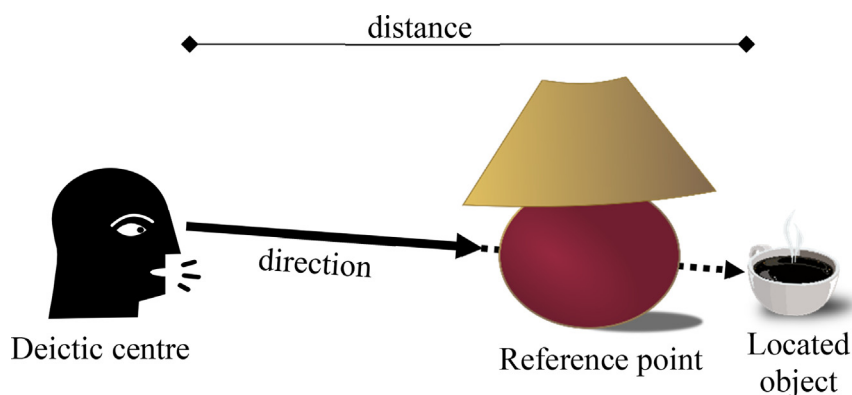


Fig. 2. “The cup is behind the lamp”.

While expressions like “behind” specify all three parameters, other expressions specify just one of them, as “near” (distance) or two, as “thither” (direction + point of reference), leaving the other(s) without any specification.

Evidential deixis can, in an analogous way, be constructed by means of three similar continuous and mutually independent parameters (direction, distance and reference points) that locate the source of information in relation to the speaker and the speech situation (Bermúdez, 2006, 2010, 2011, 2016):

1. **Mode of access to information**, a parameter that encodes the way in which the speaker gained access to the information conveyed by the utterance, and it is built as a continuum between the poles *pure perception* and *pure cognition*, with intermediate values. This parameter corresponds to *direction* in spatial deixis. Actually, it can be defined as the *direction* of the information: *towards* the speaker (as in direct perceptual impressions, where the information “hits” the speaker, so to speak) or *from* the speaker (as a result of a cognitive process inside the speaker, as in the case of inference or reasoning), with in-between values, as in the case of dreams or premonitions.
2. **Distance to the source of information**, a parameter that covers the continuum from first-hand information (the minimal possible distance), over second-hand information to hearsay and rumor (maximal distance).
3. **Source accessibility**, which describes a continuum from exclusive to universal access to the information, with intermediate values, such as shared access between speaker and hearer. Put differently, this parameter serves to indicate the evidential “referent points”, i.e. who the speaker claims has access to the information.³

Fig. 3 shows these parameters schematically:

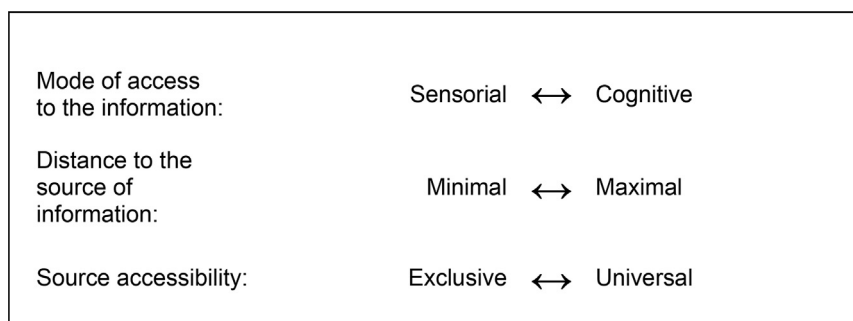


Fig. 3. Evidentiality as a deictic semantic category.

For several reasons, this model has proven to be conceptually and empirically more accurate for describing the semantic space of evidentiality than other proposed models (Willett, 1988; Frawley, 1992; Plungian, 2001, 2010; Aikhenvald, 2004; Izquierdo Alegría, 2016). On the one hand, evidential meanings such as folklore, mirativity, shared information and endophoric evidence (Tournadre, 1996) find a natural place in the system, whereas the other proposals struggle to accommodate these values within their classifications. Folklore, for example, is grouped in Willett’s model along with second-hand and third-hand evidence, i.e. hearsay, because it is neither direct nor inferring evidence, but it is clear that a central part of its meaning is lost in saying that folklore is a sort of “reported evidence”. In our model, folklore can be straightforwardly defined as *universal access* to the information + *external source*, which captures much better the essence of the concept: knowledge that anyone within a specific language community has

³ This dimension relates to what Evans et al. (2017) and Bergqvist and Kittilä (2020) call *engagement system*.

access to, but that at the same time is not first hand. Shared information between speaker and hearer (a value within the *engagement system*, see Evans et al., 2017), has no place in Willett's model, but in our deictic model it is characterized as an intermediate value between *exclusive* and *universal access* to the information, while endophoric evidence is defined as an intermediate value between *sensorial* and *cognitive*.⁴ Finally, mirativity in the sense of "sudden realization" or "unprepared mind" (Lau and Roorick, 2017) can be described in our model as a combination of speaker's *lack of access* and *cognitive* mode of access.

Likewise, compound evidential meanings that can be found in many evidential markers, such as *second-hand inference*, present in evidential adverbs such as Spanish "teóricamente" ("in theory"), or *shared inference*, as in "evidentemente" ("evidently") or "obviamente" ("obviously"), are impossible to describe in models like Willett's, because an evidential must be either a second-hand marker or an inference marker. In our model, however, these compound meanings can straightforwardly be accounted for, since these values (inference, second-hand and shared knowledge) relate to different and mutually independent parameters. Likewise the model can account for on-the-fly evidential expressions like "as everyone here has heard" (*external source + shared access*) or "as you can understand" (*cognitive mode + shared access*).

Regarding the evidential value of quotative, which is the focus of the present study, it is absent in Willett's original classification, but it would find its place along second and third hand knowledge, i.e. as reported indirect evidence. In fact, the quotative is often described in terms of its relation to hearsay, not seldom attached to the discussion of the boundaries between evidentiality and reported speech (Mushin, 2001; Chojnicka, 2012; Cruschina and Remberger, 2008; Mañoso-Pacheco 2019 among others). Aikhenvald (2004), for example, states that quotative involves an overt reference to the quoted source whereas hearsay evidentiality do not. In our model, the difference between second-hand evidence and quotative resides in that quotative specifies not just external source but also sensorial mode of access to knowledge.

4. Materials and methods

As mentioned before, the first question the present work attempts to answer is if there are such things as "prosodic evidential markers", that is, prosodic patterns that by themselves mark a statement with an evidential value, regardless of the lexical material to which they are attached. In order to investigate whether prosodic signals are used to convey evidential values, a Discourse Completion Task (DCT) was carried out with 15 Spanish native speakers between 24 and 62 years old, six males and nine females. All had the same regional variety, namely River-Plate Spanish. They were presented with five situations, each of them with an associated sentence to utter. The sentences were very simple, such as "Él tiene problemas con su mujer" ("He has problems with his wife") or "David subió para hablar con los vecinos de arriba" ("David went upstairs to talk to the neighbors"). The informants were asked to read the situation, trying to really get into it, at the same time as the interviewer read it aloud, and then they were prompted to utter a sentence directly related to the situation. An English translation of one of the five situations is shown here:

- (1) A friend comes to visit but David, your roommate, is not at home. He has gone upstairs to talk to your neighbors, who are also your friends, about going away together for the holidays. Your friend asks, "Where is David?" to which you answer:
David went upstairs to talk to the neighbors.

Later on, the informants were presented with new situations and they had to utter the very same sentences, with the only difference that a word or a phrase had quotation marks, as "Él tiene 'problemas' con su mujer" ("He has 'problems' with his wife") or "David subió para 'hablar' con los vecinos de arriba" ("David went upstairs to 'talk' to the neighbors"). Here is a translation of the new situation related to the previous example:

- (2) There's a loud party upstairs. David, your roommate, is very annoyed with the noise and he says to you "I'm going upstairs to talk to them". Just as he leaves the apartment and goes upstairs, the elevator door opens, it's a friend of yours coming to visit. Your friend sees David as he walks up the stairs visibly irritated. Then your friend approaches you (you are at the door of your apartment) and he asks you what is going on with David. You say:
He went upstairs to "talk" to the neighbors.

From the description of the situation in (2) it is clear that the words within quotation marks were direct quotations of another speaker (David in this case), since the very description of the situation states that David literally said "I'm going upstairs to talk to them".

The aim of the DCT was to obtain a large set of sentence pairs with identical lexical material and uttered by the same speakers, the only difference being the presence or absence of quotation in them, which is impossible to obtain using spontaneous conversations. This methodological decision entails, of course, some limitations concerning the nature of the data, since the sentences uttered are not spontaneous speech. A specific problem using DCT, shared by other studies that use this method of collecting data and linked to the "non-authentic" character of the collected utterances, is that the informants,

⁴ Dreams, which use to be marked by means of the choice of tense (imperfect and pluperfect) in Spanish, can also be characterized as a value between sensorial and cognitive:

- i) Yo estaba en una ciudad que era Buenos Aires pero al mismo tiempo estaba en Europa, y de repente me atacaban unos niños que de alguna manera habían adquirido poderes sobrenaturales. '(I dreamed that) I was in a city that was Buenos Aires but was in Europe at the same time, and suddenly I was attacked by some kids who had somehow gained superpowers.'

as they are told to pretend to be a character in the presented situation, can deviate from their “normal” voice in order to animate the character they pretend to be. This could certainly be a problem if the aim of this study was to establish global characteristics of their speech. However, any eventual differences between the informant’s unquoted “normal” voice and their unquoted “in-character” voice are not of relevance here, as the present study focuses only on the differences between the quoted snippet and the rest of the utterance, i.e. how the speaker *marks* the quotation, regardless of how their “normal” voice really sounds like. Nevertheless, it would be interesting to collect additional data in less controlled environments, for example from everyday conversations, to confirm the results of this study.

Another problem with the design of the study is that the situations attached to the utterance with the quotation, as in (2) above, are such that the informants not merely quote another person’s discourse but also distance themselves from it, adding some irony, as can be noticed in (2).⁵ The possible effects of this will be discussed later.

It is worth to discuss briefly the relationship between reported speech and evidentiality, not merely taking the evidential character of the example in (2) for granted. The main purpose of reported speech is to attribute a fragment of discourse to an identified person, evoking the idea of a preceding speech act to which it refers. In other words, the goal of reported speech is merely to communicate that some individual has said something. Unlike reported speech, evidentiality is strongly linked to epistemicity (Boye, 2012) and the very notion of knowledge, which in the standard analysis (Williams, 2001) is defined as *justified true belief* (Gettier, 1963). Evidentiality is connected to the *justification* part of that definition: by using evidentials, speakers justify conveying certain information by referring to the manner in which such information was accessed, the source that generated it, or the individuals who have access to it (see Fig. 3 above). More concretely, a quotative evidential justifies the use of a bit of information by evoking the source from which that information originated, signalling that the speaker had only indirect access to it (see Vanderbiesen, 2014 for a discussion). In other words, quotative evidentials and reported speech, although related, are different linguistic phenomena, that have different communicative goals and act in different ways.

This can be seen in the contrast between, on the one hand, (3)–(5), where the sole purpose of the speaker is to attribute a fragment of discourse to David, and by no means to justify some of the speaker’s beliefs or speech acts, and on the other hand (6) (and (2) above), where the speaker justifies conveying the information that David went upstairs to talk to the neighbors by evoking its source, in this case the same David.

- (3) David said: “I’m going upstairs to talk to the neighbors.”
- (4) ...and David goes “I’m going upstairs to talk to the neighbors”.
- (5) David said that he would go upstairs to talk to the neighbors.
- (6) According to himself, David went upstairs to talk to the neighbors

The examples collected by the DCT were further phonetically analyzed using the computer program PRAAT, looking for prosodic patterns. This could tell if there were some prosodic cues that mark the quotation as such, i.e. prosodic traits that work as the quotation marks in the written examples. An example of a sentence analyzed by the program can be seen in Fig. 4:

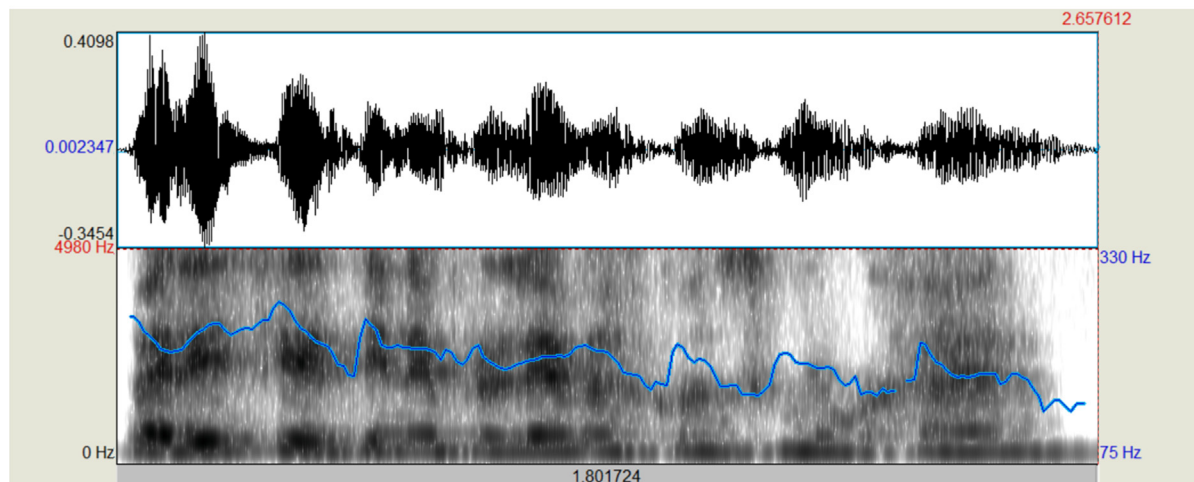


Fig. 4. Example of a sentence analyzed with PRAAT.

⁵ I assume here a definition of verbal irony that includes but is not restricted to antiphrasis, i.e. saying the opposite of what is really meant, along the lines of Abrams and Harpham (2009): “Verbal irony is a statement in which the meaning that a speaker employs is sharply different from the meaning that is ostensibly expressed, in a way that makes it obvious what the true intention is.” Concretely, I take the position of Gibbs (2000), who suggested that tropes like jocular, hyperbole and understatement should be included in the category of verbal irony. One advantage of taking this position is that other works that study the prosodics of irony, as Bryant (2010), which I compare my results with, take the same approach.

The upper section of the chart shows the intensity of the sound, expressed in Pascal, which can be described as variation in loudness level over time. The lower part of the chart is the spectrogram, which depicts the frequency of the sound, and the continuous blue line in the spectrogram is f_0 , the fundamental frequency, which is experienced as pitch. The PRAAT program makes it possible to automatically calculate, for a given fragment, both the maximum and minimum peaks of intensity and tone, as well as their mean values. The chart in Fig. 4 shows a typical assertive sentence in Spanish, with a gradually descending f_0 and intensity from the first stressed syllable to the final cadence.

5. Results

The results showed that all of the informants consistently used prosodic cues to mark the quoted word(s) in their utterances. These marks were

- a pause before the quoted expression, and often even after it.
- a raise of f_0 , i.e. a higher pitch in the quoted fragment.
- lower speed

An additional prosodic trait that was found in most (but not all) cases was a greater amplitude of the quoted discourse, i.e. a higher volume. These prosodic markers can be observed by comparing Fig. 5 with Fig. 6:

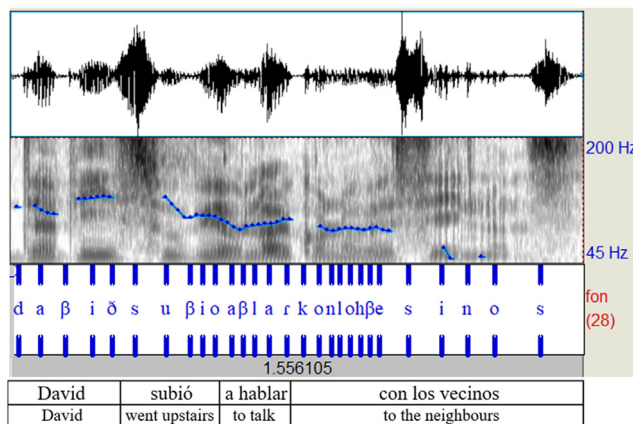


Fig. 5. David went upstairs to talk with the neighbors.

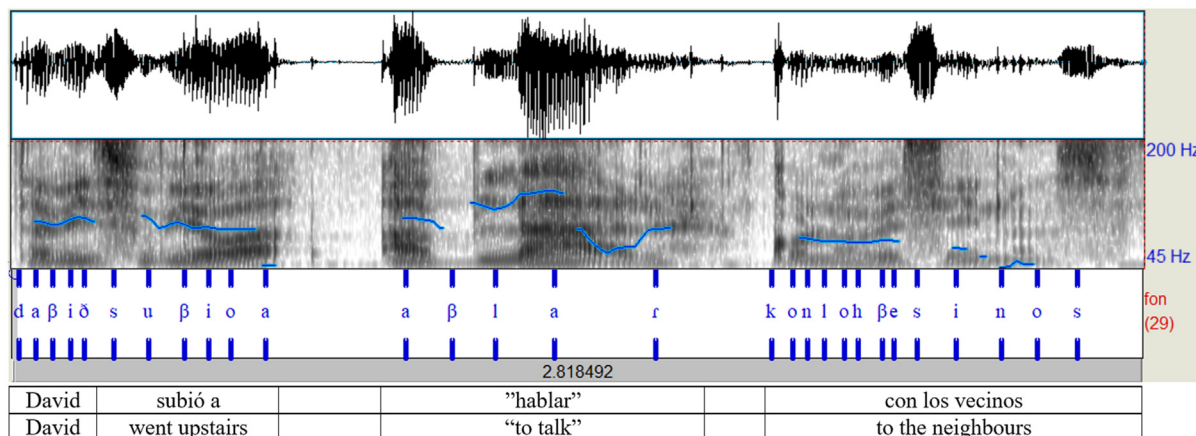


Fig. 6. David went upstairs to “talk” to the neighbors.

In comparing Figs. 5 and 6, all the mentioned traits appear clearly: in Fig. 6, the one with the quotation, a distinct pause can be noticed before and after “hablar” (the “whiter” areas in the spectrogram, which last 0.21 and 0.11 s respectively), as well as a raise of the pitch in the quote (depicted by the blue line), a higher volume (the intensity indicator above the spectrogram) and a slower tempo around the stressed syllable (evident in the relative length of the word “hablar”, and in particular its second syllable). An interesting detail is that in the sentence “subió a ‘hablar’ con los vecinos” there is a hiatus between the two “a” of the phrase “a hablar” (“to talk”), which does not happen in cases without quotation. The standard solution for this configuration of two equal unstressed vowels in Spanish is to reduce them to a single vowel [a.βlar]. However, the break between the preposition “a” and the quote “hablar” creates a clear hiatus: [a.aβlar].

Even if the differences can be clearly spotted in the diagrams, a comparison of the numerical values of the parameters just mentioned in both sentences can further clarify their significance. Regarding pitch contour, we can see that the melodic pattern of the sentence without quotation fits the typical melodic structure of simple assertions in Spanish: a raise of the pitch up to the first stressed syllable with a consistent decreasing pattern down to the end of the sentence. Table 1 shows this pattern in absolute numerical values:

Table 1
Average pitch and pitch peaks in the sentences with and without quotation.

	Sentence without quotation		Sentence with quotation	
	Average pitch	Pitch peak	Average pitch	Pitch peak
David subió a	111.71 Hz	126.87 Hz	97.85 Hz	106.44 Hz
hablar	93.21 Hz	98.98 Hz	112.65 Hz	138.84 Hz
con los vecinos	84.40 Hz	90.26 Hz	73.33 Hz	81.40 Hz

In the sentence without quotation, corresponding to Fig. 5, the average pitch decreases constantly through the sentence, the highest peak being coincident with the first stressed syllable, which is the second syllable of the word “David”. On the contrary, in the sentence with quotation, the picture is exactly the opposite: the fragment that constitutes the quotation (“hablar”) has a higher average pitch (112.65 Hz) than “David subió a” (97.85 Hz) and also a much higher peak (138.84 against 106.44 Hz), which coincides with the stressed syllable of “hablar”.

Table 2
Average intensity and intensity peaks in the sentences with and without quotation.

	Sentence without quotation		Sentence with quotation	
	Average intensity	Intensity peak	Average intensity	Intensity peak
David subió a	63.78 dB	69.92 dB	63.59 dB	67.27 dB
hablar	63.60 dB	67.55 dB	65.24 dB	70.07 dB
con los vecinos	60.97 dB	67.32 dB	58.15 dB	66.25 dB

A similar picture emerges regarding intensity. Table 2 shows that intensity slightly decreases throughout the sentence without quotation, both in average and regarding the peaks. On the contrary, the quoted fragment has both a higher average intensity than the rest of the sentence and also the highest intensity peak in the whole sentence.

Finally, Table 3 shows an analogous pattern regarding tempo:

Table 3
Relative length of *hablar* in the sentences with and without quotation.

	Sentence without quotation	Sentence with quotation
David subió a	0.60 s	0.68 s
hablar	0.17 s	0.81 s
con los vecinos	0.77 s	0.93 s

Whereas the quoted word “hablar” lasts 0.81 s (which constitutes 33% of the length of the whole sentence), the same word lasts only 0.17 s in the sentence where “hablar” is not a quotation (only 11% of the whole sentence). It is worth noting that it is in the stressed syllable of the quoted word “hablar” ([βlar]) where the decreasing of tempo is most striking, as it lasts 0.66 s.

A degree of variation was noted in the intensity of these prosodic traits: longer or shorter pauses, bigger or smaller pitch changes, etc., but all of them were present in each quotation.

It is worth pointing out that the traits found in our results to some extent resemble those of the studies that investigate the prosodics of reported speech referred to in Section 2. This is not surprising, as reported speech and quotative evidentiality, even if they have different communicative goals, as we claimed in Section 4, crucially share the reference to a previous speech act.

However, a question arises from these results: Are these features characteristic for Spanish (or even for this specific dialect of Spanish) or do they also occur in other languages?

In order to examine if the prosodic patterns found in Spanish are an idiosyncratic trait of River-Plate Spanish or if they are a more widespread phenomenon, an extension of the original study was conducted. The DCT of the original study was translated into three languages from three different language families, namely Polish (Slavic), Swedish (Germanic) and Hungarian (Finno-Ugric). Two informants from each language underwent the same study as the original study in Spanish.

The results were as conclusive as in the Spanish study: despite some variation in the intensity of the traits, all informants show the same pattern, namely a pause before the quoted expression and often even after it, a raise of f0 of the quote, a slowdown of the tempo and, often but not always, an increase of the intensity of the quoted word or phrase. We will illustrate the results by presenting one sentence pair from each language and briefly commenting on it. Figs. 7 and 8 represent the Polish examples.

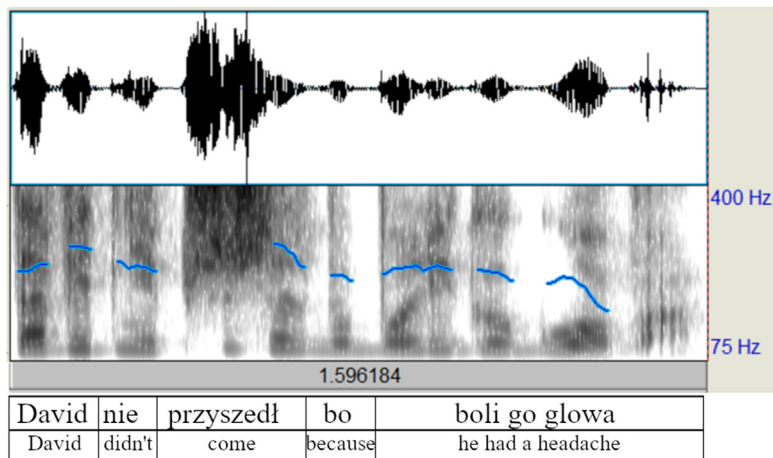


Fig. 7. David didn't come because he had a headache (Polish).

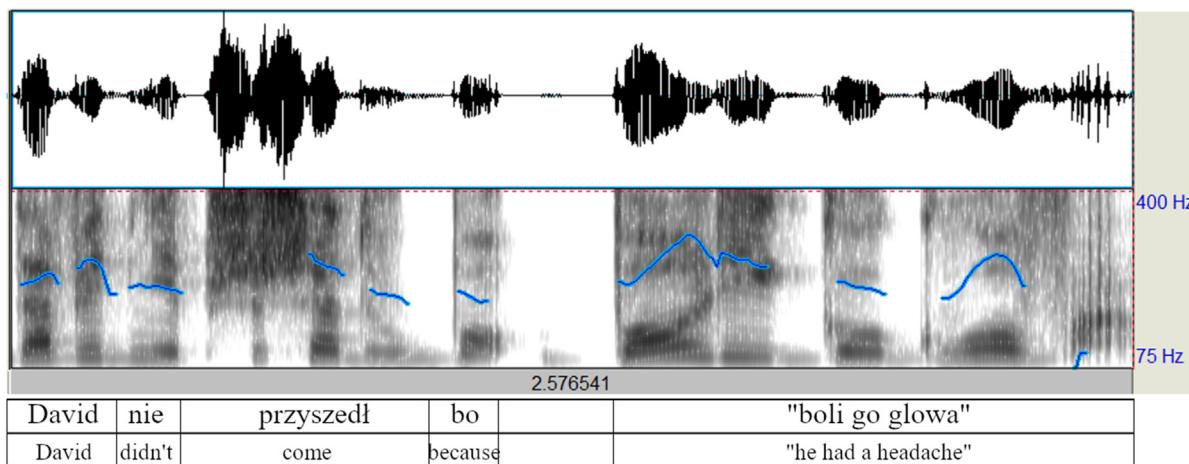


Fig. 8. David didn't come because "he had a headache" (Polish).

Again, the main traits can be seen graphically in the diagrams, but an exact measure of them can give a more accurate picture of the phenomenon. First, the pause before the quotation is 0.27 s, which is a very long pause taking into account that the informants were actually reading the sentence aloud and therefore they do not need time to think about the wording, for example. Here, of course, there is no pause after the quotation because the quotation is at the end of the sentence.

Table 4
Polish. Average pitch and pitch peaks in the sentences with and without quotation.

	Sentence without quotation		Sentence with quotation	
	Average pitch	Pitch peak	Average pitch	Pitch peak
David nie przyszedł bo	256.26 Hz	290.41 Hz	228.85 Hz	272.50 Hz
boli go głowa	229.69 Hz	249.72 Hz	255.82 Hz	317.59 Hz

Table 4 shows a similar pattern as the Spanish example: in the sentence without quotation, the phrase “Boli go glowa” (“he had a headache”) has both a lower average pitch and a lower peak than the rest of the sentence, which is consistent with the non-marked intonation pattern of assertions in Polish, as the phrase appears at the end of the sentence. On the contrary, the same phrase, when it constitutes a quotation, has a higher average pitch than the rest of the sentence and the highest peak in the whole sentence.

Regarding intensity, the Polish example shows a somewhat different pattern than the Spanish example, as neither the average intensity nor the intensity peak of the quotation is higher than the rest of the sentence in absolute terms (see Table 5).

Table 5
Polish. Average intensity and intensity peaks in the sentences with and without quotation.

	Sentence without quotation		Sentence with quotation	
	Average intensity	Intensity peak	Average intensity	Intensity peak
David nie przyszedł bo	70.45 dB	78.35 dB	68.71 dB	77.21 dB
boli go glowa	64.79 dB	71.77 dB	68.56 dB	76.06 dB

Nevertheless, it is worth noting that the values are higher than one would expect considering that the phrase “boli go glowa” is placed at the very end of the sentence, where the intensity is expected to decay. This fact becomes clear by comparing the intensity of “boli go glowa” in both sentences: the average intensity of the phrase is 6 dB lower than the preceding discourse when it is not a quotation, whereas the quoted phrase has virtually the same intensity as the rest of the sentence. Therefore, even if the absolute values of the quotation are not higher, they are nevertheless higher than expected.

Finally, the rhythm variations follow the same pattern as the Spanish example, as the phrase “boli go glowa”, which lasts exactly the same amount of time as the rest of the sentence in the example without quotation, lasts 38% more in the other sentence, due to the *rallentando* of the quotation (see Table 6)

Table 6
Polish. Relative length of *boli go glowa* in the sentences with and without quotation.

	Sentence without quotation	Sentence with quotation
David nie przyszedł bo	0.80 s	1.14 s
boli go glowa	0.79 s	1.58 s

Figs. 9 and 10 below are examples from Swedish:

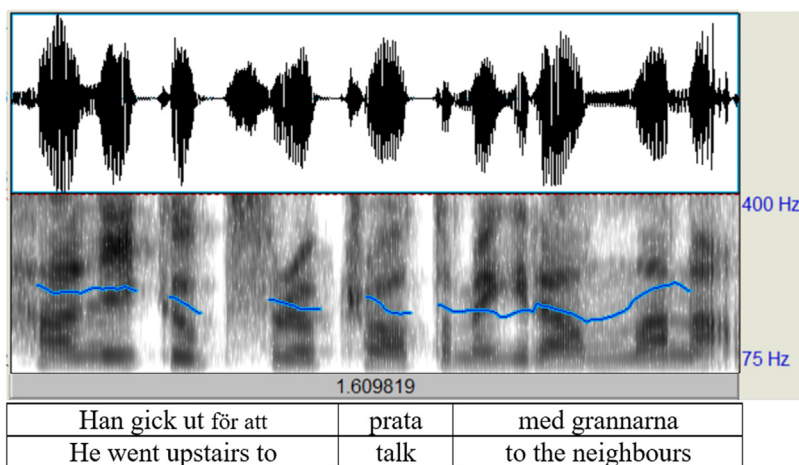


Fig. 9. He went upstairs to talk with the neighbors (Swedish).

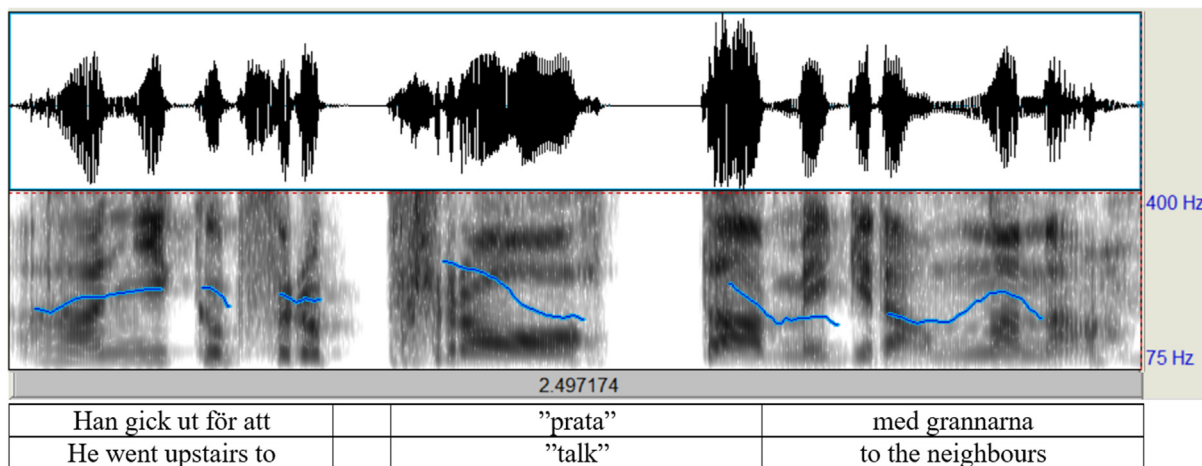


Fig. 10. He went upstairs to “talk” with the neighbors (Swedish).

Fig. 10 shows a paus of 0.14 s before the quotation “prata”. However, there is no pause after the quotation. As we pointed out earlier, all the informants made a pause before the quotation, but not always after. It would be interesting to pursue this phenomenon further: which factors can influence the emergence/absence of a pause after quotation.

A comparison between Figs. 9 and 10 shows that, again, both the average pitch and the pitch peak are highest in the quotation, as Table 7 shows in absolute numbers.

Table 7
Swedish. Average pitch and pitch peaks in the sentences with and without quotation.

	Sentence without quotation		Sentence with quotation	
	Average pitch	Pitch peak	Average pitch	Pitch peak
Han gick ut för att	212.59 Hz	232.53 Hz	205.40 Hz	223.67 Hz
prata	193.52 Hz	212.06 Hz	213.82 Hz	269.20 Hz
med grannarna	194.94 Hz	236.08 Hz	179.00 Hz	215.96 Hz

Table 8 shows further how the word “prata” (“talks”) is highlighted with a higher intensity when it constitutes a quotation, which becomes clear as the same word has a clearly lower intensity than the rest of the sentence when it is not a quotation, which can be seen graphically by comparing Figs. 9 and 10 above.

Table 8
Swedish. Average intensity and intensity peaks in the sentences with and without quotation.

	Sentence without quotation		Sentence with quotation	
	Average intensity	Intensity peak	Average intensity	Intensity peak
Han gick upp för att	72.02 dB	77.55 dB	73.25 dB	79.27 dB
prata	69.36 dB	75.36 dB	76.29 dB	81.97 dB
med grannarna	70.57 dB	75.62 dB	70.90 dB	76.29 dB

Finally, Table 9 shows that the word “prata” has a lower tempo when it is a quotation: when the word “prata” is not a quote, the word lasts approximately half as long as what precedes and follows it (0.29 s versus 0.65 and 0.62 s, respectively), while when the word is a quote, it lasts longer than both the preceding and subsequent speech (0.81 versus 0.68 and 0.71, respectively).

Table 9
Swedish. Relative length of *prata* in the sentences with and without quotation.

	Sentence without quotation	Sentence with quotation
Han gick upp för att	0.65 s	0.68 s
prata	0.29 s	0.81 s
med grannarna	0.62 s	0.71 s

Finally, Figs. 11 and 12 are examples of Hungarian informants:

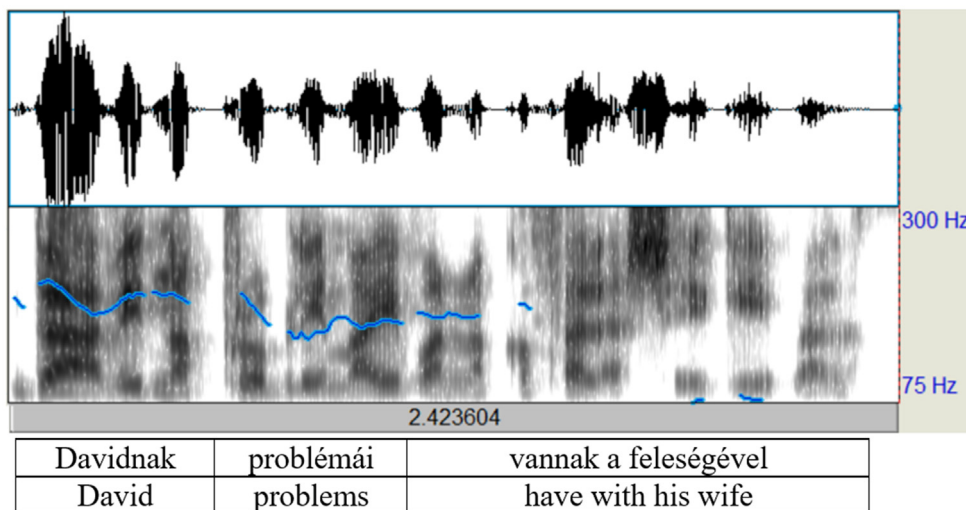


Fig. 11. David has problems with his wife (Hungarian).

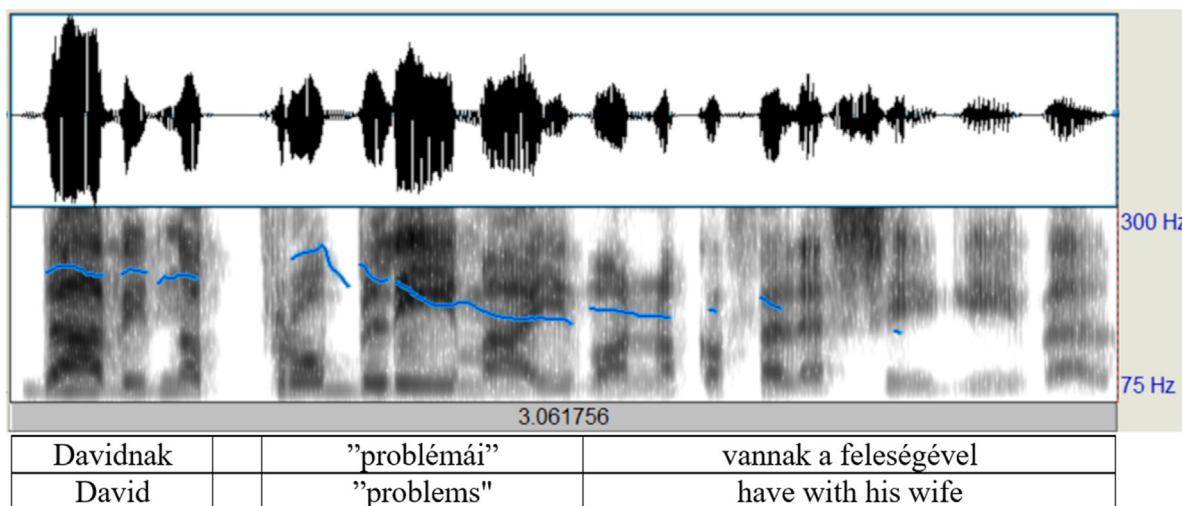


Fig. 12. David has "problems" with his wife (Hungarian).

Like in the Polish example, the informant makes a pause before the quotation (the white area before the word "problémái" in Fig. 12, that lasts 0.12 s) but there is not a clear audible pause after, even if the spectrogram shows a slight white area after the word "problémái".

Regarding the pitch, an interesting picture arises. As Table 10 shows, the average pitch of the quotation is in fact lower than the fragment before (205.33 Hz versus 220.23 Hz). This seems to contradict our claim that informants in all four languages mark the quotation with a higher pitch. Nevertheless, it is necessary to interpret these numbers in their context. Even if the average pitch of the quotation "problémái" is lower than the phrase before "Davidnak", we claim that its average pitch is still higher than expected and therefore it is experienced as higher. Firstly, if we compare these numbers with those of the sentence without quotation, the decreasing of the pitch is much smaller (-15 Hz versus -26 Hz when "problémái" do not constitute a quotation) and after the quotation the pitch goes down dramatically (205.33 Hz–159.31 Hz), whereas in the

sentence without quotation the pitch actually raises (168.33–177.65). Secondly, and even more important, the pitch peak associated with the quotation is the highest in the whole sentence (255.81 Hz). This produces an overall impression of a higher pitch.⁶

Table 10
Hungarian. Average pitch and pitch peaks in the sentences with and without quotation.

	Sentence without quotation		Sentence with quotation	
	Average pitch	Pitch peak	Average pitch	Pitch peak
Davidnak	194.58 Hz	215.25 Hz	221.23 Hz	232.21 Hz
problémái	168.33 Hz	199.87 Hz	205.33 Hz	255.81 Hz
vannak a feleségével	177.65 Hz	189.15 Hz	159.31 Hz	195.00 Hz

Regarding intensity, the picture is rather inconclusive. Table 11 shows that when “problémái” constitutes a quotation, both the average intensity and the intensity peak of the phrase are closer to the top intensity of the sentence than when “problémái” is not a quotation, but these numbers are not clear enough to claim that a listener would experience the quotation as louder. That was why we stated before that greater amplitude was found in the quoted discourse in most (but not all) cases.

Table 11
Hungarian. Average intensity and intensity peaks in the sentences with and without quotation.

	Sentence without quotation		Sentence with quotation	
	Average intensity	Intensity peak	Average intensity	Intensity peak
Davidnak	76.50 dB	82.22 dB	79.37 dB	85.29 dB
problémái	70.61 dB	76.38 dB	75.83 dB	81.83 dB
vannak a feleségével	67.22 dB	74.66 dB	67.31 dB	75.61 dB

Finally, Table 12 shows that, like in the other three languages, there is a slower tempo in the quotation, as the word “problémái”, which lasts exactly the same amount of time as the preceding phrase “Davidnak” when it is not a quotation (0.53 s versus 0.51 s), becomes 66% longer when it is a quotation (0.85 s).

Table 12
Hungarian. Relative length of “problémái” in the sentences with and without quotation.

	Sentence without quotation	Sentence with quotation
Davidnak	0.51 s	0.51 s
problémái	0.53 s	0.85 s

These results by no means imply that these prosodic traits constitute a universal marker of quotation, given the size of the study and the fact that only four languages were investigated. Nevertheless, the results suggest that the first question of this study must have an affirmative answer, i. e., that the study has shown that there is a (rather widespread) pure prosodic marker of quotation, that seems to consist of a cluster of prosodic traits, namely intonational (pitch shift), rhythmic (pause and rallentando) and in most but not all cases, dynamic (volume increase).⁷

6. The iconic nature of language

What about the second question asked at the beginning, the one about the arbitrariness or motivation behind the specific traits of the prosodic markers? Let us take a closer look at the characteristics of the prosodic marking of quotation found in the study.

The pause before (and after) the quotation is clearly a signal that something is unusual, an instruction for the hearer that the snippet needs further interpretation. In this case, that the fragment that follows is a verbatim quotation of another speaker. However, is it an arbitrary signal? We will argue that it is not.

⁶ Another factor that contributes to the impression of a higher pitch is the range of pitch variation. The maximum and minimum pitch of ‘problémái’ when it is marked as a quotation are 255.81 Hz and 165.44 Hz respectively, i.e. a range of 90 Hz. The corresponding values of ‘problémái’ when it is not marked as a quotation are 199.87 Hz and 149.26, i.e. a range of just 50 Hz.

⁷ Later we will call this statement into question, as we argue that not all of the traits are linked to the quotative value.

A pause⁸ is something that prototypically occurs in a conversation when the speaker changes, for example, between a question and the subsequent answer. Of course, speaker change in normal conversation can occur with no gap, and speakers can overlap each other, but a pause is the prototypical, unmarked situation (Heldner and Edlund, 2010). In fact, both overlap and the lack of a break between two conversational turns are marked in Jefferson's (1984, 2004) transcription conventions (with [] and =, respectively), but a brake between two conversational turns is not marked, which is evidence of its prototypical nature. Therefore, a pause in the middle of a sentence, when it is not expected, is a given candidate for marking a fragment as a quotation, since a quotation entails a metaphorical shift of speaker: it is a fragment of discourse attributed to another individual. In other words, there seems to be a motivation for this specific prosodic trait to have become a marker of quotation. The opposite strategy, for example to accelerate the tempo of the last word before the quotation and the onset of the quotation itself to be sure that no pause arises between the words would be a much worse choice to mark this metaphorical speaker shift. The fact that languages from different language families adopt the same strategy is additional evidence for the claim that this phenomenon is one that is deeply embedded in human communication patterns rather than a peculiar characteristic of a given language.

We will argue that the second trait observed, i.e. a pitch raise in the quoted fragment, is motivated too. As a matter of fact, the most identifiable aspect when the speaker changes in a conversation is the change in voice quality: they are after all two different speakers with two different voices. To a “blind” observer, who only has access to prosodic cues, a change in voice quality would be the most straightforward indication that a new speaker has taken the floor. Changing the pitch is then a simple and straightforward way to change voice quality, suggesting a change of enunciator; in other words, a metaphorical change of speaker. Therefore, we argue that the choice of this specific trait is not random but motivated. However, by changing the pitch of a fragment the speaker is not trying to, with minimal effort, mimic the voice of the specific quoted individual, as one might expect. If this were the case, the speaker would either raise or lower the pitch depending on the quoted voice, but in the study, all the informants actually raised the pitch, even when the informant was a woman and the quoted individual was a man. In other words, the change in voice quality is not just mimicry.

It is worth wondering why all speakers in the four languages used a higher pitch. In principle, if it had been just a matter of convention, a lowering of the pitch could likewise have become (part of) the standard marking of speaker change, instead of pitch raising, if the goal was merely to suggest a new enunciator by changing voice quality. In other words, why has pitch raising become part of the marking of quotation rather than, for example, pitch lowering? We will advance a rather speculative hypothesis to give a tentative answer to the question, or at least a line of inquiry, that connects with the ecological aspect of language, i.e. language embodiment, one of the pillars of cognitive linguistics (Lakoff, 1987, 2012; Lakoff and Johnson, 1980, 1999; Kreiner and Eviatar, 2014; Esposito and Gratton, 2022).

When people are told to look at themselves, they tend to lower their heads and look down at their bodies. Conversely, when people focus on something other than themselves they tend to lift their head. This association between self-focusing and looking down is also true in a more abstract sense: for example, a picture with a man looking down can convey self-oriented personality traits like introversion or shyness, whereas a picture with a man with his head lifted, looking forward, can convey personality traits like extroversion or sociability, but arguably not the other way around. This is also the case regarding linguistic expressions: put to choose, discourse markers such as “Don't you think?” would be associated with the picture with the man with his head up while those like “I think” would be associated with the one with his head down, and not vice versa. At the same time, anatomical factors favor higher pitch when the head is lifted and lower pitch when the head is facing down. If this is the case, a correlation between markers of hearer-oriented speech acts and higher pitch or contour could probably be found, like between markers of more speaker-oriented speech acts and lower pitch. Empirical data can corroborate or challenge this hypothesis; nevertheless, there is some data that points in that direction. There are several examples of high tones or contours related to hearer-oriented functions, like the intonation of questions or the intonation of statements for which the speaker claims shared knowledge, whereas lowering of the overall pitch tends to relate to complete different situations, for example portions of discourse about which the speaker closes the possibility of discussion with the hearer, as in explicative relative clauses, see (7) below:

(7) The explanation, *which is unquestionable*, is that the president has lied to the people.

In accordance with the discussion above, it would be counterintuitive to associate quoting someone else's discourse with downward head motion and lower pitch, since quoting someone else's discourse entails focusing on other than oneself.⁹ This could be argued as a possible line of explanation for pitch raising becoming (part of) the marking of quotation rather than pitch lowering.

To sum up the discussion so far, the proposal is that speakers use pause and change in voice quality (specifically, pitch raising to convey that a change of enunciator has occurred. These markers seem to be motivated rather than random, as these features typically appear in a conversation when the speaker changes: they would actually be the two main signals to

⁸ It is usual to label intervals within the speech of one speaker as *pauses* and intervals between speakers or at speaker changes as *gaps*. Here we use the terms *pauses*, *gaps* or *breaks* indistinctly.

⁹ Additional evidence in favour of this interpretation comes from work in kinetics and head motion in particular. McClave (2000) find that “a speaker's head often will assume a new orientation slightly preceding or coinciding with the beginning of a quote”; nevertheless, this new orientation is never downwards, but to the side or up.

a blind observer that a new speaker has taken the floor. In other words, they have to be interpreted iconically, as a marker of a metaphorical change of speaker; they seem to be “iconically logical” traits to function as markers of quotation. It can be said that speakers use these traits metonymically, as they are prosodic clues of speaker change.

The remaining two traits observed, namely slower tempo around the stressed syllable and higher intensity of the quotation, could possibly be attributed to the same goal, i.e. a marking of the change of the enunciator, by means of changing some aspect of the voice quality. Nevertheless, we will here pursue an alternative explanation. It is not obvious that all of the traits observed actually function as prosodic marking of quotation; some of them could be related to other meanings also conveyed by the quotation. As stated earlier, the design of the study implied that the informants not only conveyed that a given expression was a quotation but they also distanced themselves from it, reproducing the quoted expression with some irony. When they for example said that David went upstairs to “talk” to the neighbors, they did not only mean that the word “talk” is a quotation of David’s discourse, but also that in their interpretation of the situation, David did not just go upstairs to talk to the neighbors but rather to *complain* to them. Could the change in speed and the higher intensity relate to the irony rather than to mere quotation? Results of other unrelated studies suggest that this can be the case. Skrelin et al. (2020: 552), for example, find precisely these two characteristics in their study about irony in Russian: “The acoustic analysis of ironic snippets ... showed that ironic stimuli are characterized by a higher intensity of the target fragment and a longer duration of the stressed vowel.” Similar results are found for English (Rockwell, 2000), German (Scharrer and Christmann, 2011), French (Deliens et al., 2018) and Spanish (Rao et al., 2022). Bryant (2010), in his study of ironic speech in English conversations, finds that even if a higher intensity was found in many informants, the only trait that was consistent across all the informants was that ironic utterances were spoken significantly slower than preceding speech. This coincides with our results: a lower speed was consistently used by all informants whereas higher intensity was found in most cases but not all of them.

Given that these two prosodic characteristics were the ones found as the only two specific prosodic markers for irony and that irony is a component of the sentences uttered by our informants, it is not unthinkable that these two prosodic traits in the quoted snippets belong to the expression of irony rather than to the marking of the fragment as a quotation. This hypothesis requires of course empirical evidence; an extension of the present study based on speech recognition tasks to test this hypothesis is ongoing.

7. Conclusions

The object of this study was to investigate if pure prosodic marking of quotation does exist and, secondly, if the particular prosodic traits involved in such prosodic marking are arbitrary or if some iconic relation between them and the meaning they convey could be found.

The results of a study carried out in four languages belonging to four different families (Spanish, Polish, Swedish and Hungarian) suggested that an affirmative answer must be given to the first question, as the quoted fragments were consistently marked in the four languages with a pause before the quoted fragment (and often after it) and a raise of the pitch of the snippet. Two other prosodic traits (higher intensity and slower tempo) were also found, but these traits were interpreted as conveying irony rather than quotative evidentiality, as irony was present in these quotations and precisely these two prosodic characteristics were found in unrelated studies as specific markers of irony. In order to confirm this hypothesis, a further study is needed where only quotation –and no distancing or irony– is involved.

Regarding the question of the arbitrariness or motivation of the prosodic markers, the results suggest that both of the prosodic traits found are indeed motivated. The pause found in all of the informants is a trait that appears prototypically in everyday conversation between turns, i.e. when the speaker changes. Since a quotation is a fragment of discourse that is assigned to an individual other than the actual speaker or, in other words, a change of enunciator, a pause is then a “logical” marker for it, since a pause is what normally occurs when the speaker changes. The other prosodic phenomenon in everyday conversation that for a “blind” observer is an undeniable clue of speaker change is that the quality of the voice changes, just because different speakers have different voices. The pitch shift in the quoted fragments observed in the study then works as a conventionalized change in voice quality, which signals that the fragment must be attributed to an enunciator other than the speaker. The consistent choice of pitch raising to mark the change of enunciator (as opposed to pitch lowering for example) is argued to be related to issues of language embodiment.

Much work is yet to be done but the results of this study are promising first steps in explaining the relation between prosody and evidentiality.

Declaration of competing interest

None

Data availability

Data will be made available on request.

References

- Abrams, M.H., Harpham, Geoffrey, 2009. *A Glossary of Literary Terms*. Heinle & Heinle, Boston.
- Aikhenvald, Alexandra, 2004. *Evidentiality*. Oxford University Press, Oxford.
- Aikhenvald, Alexandra, Dixon, R.M.W. (Eds.), 2014. *The Grammar of Knowledge: A Cross-Linguistic Typology*. Oxford University Press, Oxford.
- Bergqvist, Henrik, Kittilä, Seppo, 2020. Epistemic perspectives: evidentiality, egophoricity, and engagement. In: Bergqvist, H., Kittilä, S. (Eds.), *Evidentiality, Egophoricity, and Engagement*. Language Science Press, Berlin, pp. 1–21.
- Bermúdez, Fernando, 2004. La categoría evidencial del castellano: metonimia y elevación de sujeto, 22. *Bol. Lingüística*, pp. 3–31. Available at: http://saber.ucv.ve/ojs/index.php/rev_bl/article/view/1466.
- Bermúdez, Fernando, 2006. *Evidencialidad. La codificación lingüística del punto de vista*. Stockholm University, Stockholm. Available at: <http://diva-portal.org/smash/get/diva2:199511/FULLTEXT01.pdf>.
- Bermúdez, Fernando, 2010. Había sido o no había sido, he ahí la cuestión. *Pluscuamperfecto y evidencialidad en castellano*. *Studia Neophilologica* 80 (2), 203–222. <https://doi.org/10.1080/00393270802528889>.
- Bermúdez, Fernando, 2011. El pluscuamperfecto como marcador evidencial en castellano. In: Hernández Socas, E., Sinner, C., Wotjak, G. (Eds.), *Estudios de tiempo y espacio en la gramática española*. Peter Lang Publishing Group, Frankfurt am Main, pp. 43–62.
- Bermúdez, Fernando, 2016. Rumores y otros malos hábitos. El condicional evidencial del español. *Cuadernos de Lingüística* 3 (2), 35–69. <https://doi.org/10.24201/ciecm.v3i2.37>.
- Boas, Franz, 1947. Kwakiutl grammar, with a glossary of the suffixes. *Trans. Am. Phil. Soc.* 37, 203–377.
- Bolden, Galina, 2004. The quote and beyond: defining boundaries of reported speech in conversational Russian. *J. Pragmat.* 36, 1017–1118.
- Boye, Kasper, 2012. *Epistemic Meaning. A Crosslinguistic and Functional-Cognitive Study*. De Gruyter Mouton, Berlin & Boston.
- Bryant, Gregory A., 2010. Prosodic contrasts in ironic speech. *Discourse Process* 47 (7), 545–566.
- Cabedo Nebot, Adrián, Cornillie, Bert, 2018. A prosody account of (Inter)Subjective modal adverbs in Spanish. In: Guentchéva, Z. (Ed.), *Epistemic Modalities and Evidentiality in Cross-Linguistic Perspective*. de Gruyter, Berlin.
- Caldiz, Adriana, 2018. Prosody and evidentiality: about the circumflex inflection in the Spanish of Buenos Aires. In: Paper presented at 15th IPRA Conference. July 2017 Belfast. UK.
- Chafe, Wallace, Nichols, Johanna (Eds.), 1986. *Evidentiality: The Linguistic Coding of Epistemology*. Ablex, Norwood, NJ.
- Chafe, Wallace, 1986. Evidentiality in English conversation and academic writing. In: Chafe, Wallace, Nichols, Johanna (Eds.), *Evidentiality: The Linguistic Coding of Epistemology*. Ablex, Norwood, NJ, pp. 261–272.
- Chojnicka, Joanna, 2012. Reportive evidentiality and reported speech: is there a boundary? Evidence of the Latvian oblique. In: Usonienė, A., Nau, N., Dabašinskiene, I. (Eds.), *Multiple Perspectives in Linguistic Research on Baltic Languages*. Cambridge Scholars Publishing, pp. 170–192.
- Cruschina, Silvio, Remberger, Eva-Maria, 2008. Hearsay and reported speech: evidentiality in Romance. *Riv. Gramm. Gener.* 33, 99–120.
- De Haan, Ferdinand, 2005. Encoding speaker perspective: evidentials. In: Frajzyngier, Z., Hodges, A., Rood, D. (Eds.), *Linguistic Diversity and Language Theories*. John Benjamins Publishing Company, pp. 379–397.
- De Lancey, Scott, 2001. The mirative and evidentiality. *J. Pragmat.* 33 (3), 369–382.
- Deliens, Gaétane, Antoniou, Kyriakos, Clin, Elise, Ostashchenko, Ekaterina, Kissine, Mikhail, 2018. Context, facial expression and prosody in irony processing. *J. Mem. Lang.* 99, 35–48.
- Dendale, Patrick, Tasmowski, Liliane, 2001. Introduction: evidentiality and related notions. *J. Pragmat.* 33, 339–348.
- Esposito, Lewis, Gratton, Chantal, 2022. Prosody and ideologies of embodiment: variation in the use of pitch and articulation rate among fitness instructors. *Lang. Soc.* 51 (2), 211–236.
- Estellés Arguedas, María, 2015. Expressing evidentiality through prosody? Prosodic voicing in reported speech in Spanish colloquial conversations. *J. Pragmat.* 85, 138–154.
- Estellés Arguedas, María, Albelda Marco, Marta, 2014. Evidentials, politeness and prosody in Spanish: a corpus analysis. *J. Politeness Res.* 10 (1), 29–62.
- Evans, Nicholas, Bergqvist, Henrik, San Roque, Lila, 2017. The grammar of engagement I: framework and initial exemplification. *Lang. Cognit.* 10, 110–140.
- Frawley, William, 1992. *Linguistic Semantics*. Laurence Erlbaum, Hillsdale.
- García Negroni, María Marta, Caldiz, Adriana, 2016. Puntos de vista evidenciales citativos y prosodia marcada. In: Paper presented at IADA, VI coloquio de diálogos y tendencias. University of La Plata.
- Gettier, Edmund, 1963. Is justified true belief knowledge? *Analysis* 23 (6), 121–123.
- Gibbs, Raymond W., 2000. Irony in talk among friends. *Metaphor Symbol* 15, 5–27.
- González Ruiz, Ramón, Izquierdo Alegría, Rámaso, Loureda Lamas, Óscar (Eds.), 2016. *La evidencialidad en español*. Iberoamericana/Vervuert, Madrid/Frankfurt am Main.
- Guentchéva, Zlatka (Ed.), 2018. *Epistemic Modalities and Evidentiality in Cross-Linguistic Perspective*. de Gruyter, Berlin.
- Günthner, Susanne, 1999. Polyphony and the “layering of voices” in reported dialogues: an analysis of the use of prosodic devices in everyday reported speech. *J. Pragmat.* 31, 685–708.
- Heldner, Mattias, Edlund, Jens, 2010. Pauses, gaps and overlaps in conversations. *J. Phonetics* 38, 555–568.
- Izquierdo Alegría, Dámaso, 2016. *Alcances y límites de la evidencialidad*. Pamplona: Universidad de Navarra.
- Jakobson, Roman, 1957. Shifters, verbal categories and the Russian verb. In: *Selected Writings Vol 2*. Cambridge University (1971). Mouton, The Hague, pp. 130–147.
- Jefferson, Gail, 1984. On the organization of laughter in talk about troubles. In: Maxwell Atkinson, J., Heritage, J. (Eds.), *Structures of Social Action: Studies in Conversation Analysis*. Cambridge University Press, Cambridge, pp. 346–369.
- Jefferson, Gail, 2004. Glossary of transcript symbols with an introduction. In: Lerner, G. (Ed.), *Conversation Analysis. Studies from the First Generation*. John Benjamins, Amsterdam, pp. 13–31.
- Kasimir, Elke, 2008. Prosodic correlates of subclausal quotation marks. *ZAS Papers in Linguistics* 49, 67–77.
- Klewitz, Gabriele, Couper-Kuhlen, Elisabeth, 1999. Quote—unquote? The role of prosody in the contextualization of reported speech sequences. *Pragmatics* 9 (4), 459–485.
- Kreiner, Hamutal, Eviatar, Zohar, 2014. The missing link in the embodiment of syntax: prosody. *Brain Lang.* 137, 91–102.
- Lakoff, George, 1987. *Women, Fire, and Dangerous Things: what Categories Reveal about the Mind*. University of Chicago Press, Chicago.
- Lakoff, George, 2012. Explaining embodied cognition results. *Topics in Cognitive Science* 4 (4), 773–785.
- Lakoff, George, Johnson, Mark, 1980. *Metaphors We Live by*. University of Chicago Press, Chicago.
- Lakoff, George, Johnson, Mark, 1999. *Philosophy in the Flesh: the Embodied Mind and its Challenge to Western Thought*. Basic Books, New York.
- Lau, Mónica, Rooryck, Johan, 2017. Aspect, evidentiality, and mirativity. *Lingua* 186–187, 110–119.
- Lazard, Gilbert, 1999. Mirativity, evidentiality, mediativity, or other? *Linguist. Typol.* 3, 91–101.
- Mañoso Pacheco, Lidia, 2019. Reportive evidentiality. A theoretical revision. *Revista Inclusiones* 6 (4), 189–206.
- McClave, Evelyn, 2000. Linguistic functions of head movements in the context of speech. *J. Pragmat.* 32 (7), 855–878.
- Mushin, Ilana, 2001. Evidentiality and deixis in narrative retelling. *J. Pragmat.* 32 (7), 927–957.
- Nuckolls, Janis, Michael, Lev, 2014. *Evidentiality in Interaction*. John Benjamins, Amsterdam.
- Palmer, F.R., 1986. *Mood and Modality*. Cambridge University Press, Cambridge.
- Plungian, Vladimir, 2001. The place of evidentiality within the universal grammatical space. *J. Pragmat.* 33 (3), 349–357.
- Plungian, Vladimir, 2010. Types of verbal evidentiality marking: an overview. In: Diewald, G., Smirnova, E. (Eds.), *The Linguistic Realization of Evidentiality in European Languages*. De Gruyter Mouton, Berlin & New York, pp. 15–58.

- Rao, Rajiv, Ye, Ting, Butera, Brianna, 2022. The prosodic expression of sarcasm vs. sincerity by heritage speakers of Spanish. *Language* 7 (1), 17.
- Rockwell, Patricia, 2000. Lower, slower, louder: vocal cues of sarcasm. *J. Psycholinguistic* 29 (5), 483–495.
- Scharrer, Lisa, Christmann, Ursula, 2011. Voice modulations in German ironic speech. *Lang. Speech* 54, 435–465.
- Pavel, Skrelin, Uliana, Kochetkova, Evdokimova, Vera, Novoselova, Daria, 2020. Can we detect irony in speech using phonetic characteristics only? – Looking for a methodology of analysis. In: Karpov, A., Potapova, R. (Eds.), *Speech and Computer. SPECOM 2020. Lecture Notes in Computer Science*. Springer, Cham.
- Söderqvist, Erika, 2020. Informational and relational functions of evidentiality in interaction. *Stud. Neophilol.* 92 (1), 56–73.
- Squartini, Mario, 2008. Lexical vs. grammatical evidentiality in French and Italian. *Linguistics* 46 (5), 917–947.
- Squartini, Mario, 2012. Evidentiality in interaction: the concessive use of the Italian future between grammar and discourse. *J. Pragmat.* 44 (15), 2116–2128.
- Tournadre, Nicolas, 1996. Comparaison des systèmes médiatifs de quatre dialectes tibétains (tibétain central, Ladakhi, dzonkha et amdo). In: Guentchéva, Z. (Ed.), *L'énonciation Médiatisée*. Peeters, Paris, pp. 195–213.
- Vanderbiesen, Jeroen, 2014. wollen: on the verge between quotative and reportive evidential. *Yearbook of the German Cognitive Linguistics Association* 2 (1), 167–190.
- Vanrell, María del Mar, Armstrong, Meghan, Prieto, Pilar, 2017. Experimental evidence for the role of intonation in evidential marking. *Lang. Speech* 60 (2), 242–259.
- Willett, Thomas, 1988. A Cross-linguistic survey of the grammaticalization of evidentiality. *Stud. Lang.* 12, 57–91.
- Williams, Michael, 2001. *Problems of Knowledge*. Oxford University Press, Oxford.

Fernando Bermúdez is an Associate Professor of Spanish linguistics at Uppsala University, Sweden. He received his PhD in Linguistics at Stockholm University in 2004. His research interests include Evidentiality and its relation to the functional categories of Tense, Aspect and Mood, a subject on which he has published a number of articles in different journals. In 2012 he received the Thuréus Prize, awarded by the Royal Society of Sciences (Sweden) for his work in the field of evidentiality in Spanish.