

Epistemic sentence adverbs, epistemic complement-taking predicates and epistemic pragmatic markers*

Paola Pietrandrea (Lille)

Abstract

This study presents a formal and functional data-driven characterization of sentence adverbs. We define epistemic sentence adverbs as semantic predicates that validate the truth value of their scopes and are macrosyntactically dependent on them. We distinguish epistemic sentence adverbs on the one hand from epistemic complement-taking predicates – i. e., semantic predicates that validate the truth value of their scopes and govern them microsyntactically – and on the other hand from epistemic pragmatic markers – i. e., semantic predicates that validate the truth value of their scopes and that are syntactically independent of them. We also hypothesize that epistemic sentence adverbs have a specific functional role. Unlike (most) epistemic complement taking predicates, they express a non-addressable epistemic evaluation of the scope; unlike pragmatic markers they serve to qualify, rather than negotiate, the epistemic evaluation of the scope. We show that such a distinction is not necessarily inscribed at the lexical level: even though some words invariably behave as sentence adverbs, other words can be classified as epistemic sentence adverbs in some syntactic contexts, and as other epistemic markers in other syntactic contexts.

1 Introduction

In this article we present the formal and functional data-driven characterization of sentence adverbs that we formalized during the development of the annotation scheme for epistemicity applied to the MoDAL Corpus. The Modal Corpus is an openly accessible¹ trilingual resource consisting of three equivalent corpora of English, French and Italian dialogues drawn, respectively, from the Santa Barbara Corpus of Spoken American English (for English), the ESLO Corpus, the OTG Corpus and the Accueil UBS Corpus (for French), and the VoLIP

* The research presented in this paper was supported by the Maison des Sciences de l'Homme du Val de Loire (grant MoDAL 2015), the IRCOM Consortium, the “Laboratoire Liggrato de Linguistique inguistique the Groningen Meaning Bank. This paper has greatly benefited from extremely helpful theoretical discussions with Caterina Mauri, Andrea Sansò, Dylan Glynn, Bjorn Wiemer, Bert Cornillie, Tanja Mortelmans, Mario Squartini, and Patrick Dendale. An early version of this work was presented at the workshop on ‘Commitment phenomena through the study of evidential markers in Romance languages’ (2014). Thanks to Corinne Rossari, Claudia Ricci and Elena Siminiciuc, who organized this stimulating scientific event. An immense thanks to Malvina Nissim and Elisa Ghia, with whom I worked on the implementation of the annotation schema. The usual disclaimers apply.

¹ The MoDAL Corpus is accessible at the following URL: <http://modal.msh-vdl.fr/?lang=en>

Corpus (for Italian), annotated for epistemic modality. We annotated about 20,000 words per language for a total number of 2824 epistemic constructions (833 for the English Corpus, 1271 for the French Corpus, 720 for the Italian Corpus) (Pietrandrea 2018; Nissim & Pietrandrea 2017).

One of the properties annotated in the corpus was the morphosyntactic nature of the epistemic markers. The effort of developing a theoretically meaningful classification of the values describing the morphosyntactic nature of the markers led us to isolate a class of epistemic sentence adverbs that we distinguished, from both a formal and a functional point of view, from epistemic complement-taking predicates (henceforth CTPs) and epistemic pragmatic markers.

The paper is organized as follows: after providing a definition of the functional domain annotated in our corpus, i. e., epistemicity (section 2), and a definition of the linguistic units that encode it, i. e. epistemic constructions (section 3), we present in section 4 the distributional criteria that we identified to define the morphosyntactic properties of the epistemic markers. These criteria, elaborated within the formal framework of the syntax of spoken language put forward by the *Rhapsodie* group (Lacheret et al., 2019), allow theoretical room for positing an autonomous class of epistemic sentence adverbs, distinguished, on the one hand, from epistemic CTPs and, on the other hand, from epistemic pragmatic markers. On these grounds, we provide in section 5 a thorough formal characterization of epistemic sentence adverbs and, in section 6, based on the study of the annotated corpus, a functional characterization of them.

2 Epistemicity

Following Boye (2012), we characterize epistemicity as the category composed of the two subcategories of evidential justification and epistemic support. Evidential justification is the mention of the evidential source that justifies the validation of the truth of a linguistic representation (1):

1. *Stando a quanto mi hanno detto, è a casa*²
'According to what they told me, (s)he is at home'
2. *Probabilmente è a casa*
'(S)he is probably at home'

Following Pietrandrea (2018: 172) we define epistemicity as the linguistic category that “explicitly indicates the process of shared validation of a truth-value to the propositional tokens that compose a discourse”. Consistently with what has recently been proposed in the intersubjectivity framework (Simon-Vandenberg/Aijmer 2007; Mithun 2012; Rossari 2012; Kärkkäinen 2012; Traugott 2012), we include under the label of epistemicity not only the linguistic operation of monological qualification of the truth-value of a propositional content, illustrated by examples (3) through (8), but also the linguistic operation of

² Examples 1 through 10, example 13, and examples 25 and 26 are invented examples. All the other examples in this article are attested in the VoLIP component of the MoDAL corpus, to which we refer as [MoDAL-VOLIP]. The Modal Corpus is a trilingual resource which is entirely annotated now, but at the time the present article was written only the Italian component was available, which explains why all our examples are in Italian.

negotiation of the validation of a truth-value between two or more speakers, illustrated in (9) and (10).

3. *Sarà a casa.*
'(S)he will be at home.'
4. *Deve essere a casa.*
'(S)he must be at home.'
5. *A quanto pare è a casa.*
'Apparently, (s)he is at home.'
6. *E' sicuramente a casa.*
'(s)he is surely at home.'
7. *Forse è a casa.*
'Maybe (s)he is at home.'
8. *E' a casa, penso.*
'(S)he is at home, I think.'
9. A: *E' a casa.* B: *davvero?*
'A: (S)he is at home B: really?'
10. A: *E' a casa.* B: *eh si', è a casa.*
'A: (S)he is at home B: oh yes, (s)he is.'

3 Epistemic constructions

We adopted for our work a general constructional approach, which led us to claim that epistemicity is encoded by epistemic constructions, i. e. conventional associations of a form and a function, which can be identified at any level of linguistic structure and complexity.

Following Pietrandrea (2018), we represent epistemic constructions as complex constructions consisting of: (i) an epistemic marker, (ii) an epistemically modalized scope, and (iii) the relation between the epistemic marker and the epistemically modalized scope.³

11. [[[*Probabilmente*]_{epm} [*è a casa*]_{eps}]_{epr}]_{epcon}
'Probably she is at home'
12. A: [[[*Ha mangiato e ha dormito*]_{eps}
'A: (s)he ate and slept'
B: [*certo*]_{epm}]_{epr}]_{epcon}
'B: of course [lit. certain.M.SG]'

3.1 The epistemic marker

As we will see below, epistemic markers can be realized by various formal elements: morphemes, modal verbs, sentence adverbs, pragmatic markers, epistemic utterances, and

³ In Pietrandrea (2018) we show that each of these three elements can be regarded as a construction *per se*.

prosodic profiles. From a semantic standpoint, however, epistemic markers constitute a unitary category comprised of all the semantic predicates (cf. Polguère 1992; Mel'cuk 2014) that validate the truth-value of their arguments.⁴ We will provide in this section a technical definition of “semantic predicate”. For the moment, an intuitive definition is that the epistemic marker of the construction (11) is the adverb *probabilmente*, and the epistemic marker of the construction (12) is the pragmatic marker *certo*.

3.2 Epistemic scope

The scope of an epistemic construction corresponds to the element of the context whose truth-value is validated by the marker. As a truth-value bearer, the scope of an epistemic construction expresses, by definition, a proposition (or several propositions), i. e., abstract semantic objects capable of being evaluated in terms of truth.⁵ The epistemically modalized scope of the construction (11) corresponds to the proposition *è a casa*. The epistemically modalized scope of the construction (12) corresponds to the two propositions *ha mangiato* and *ha dormito*.

3.3 The epistemic relation

The relation between the marker and the scope corresponds to the acknowledgment that in a discourse a given marker is associated to a given scope. In (11) for example, a relation is established between the epistemic marker *probabilmente* and the scope *è a casa*, in (12) between the marker *certo* and the scope *ha mangiato e ha dormito*.

The epistemic relation is described formally by the mutual position of the marker and the scope, as well as by the distribution of the marker and the scope across turns of speech and across speakers. The relation in (11) can be described as a relation of precedence of the marker with respect to the scope and a co-occurrence of both in the same speech turn, whereas the relation in (12) can be described as a relation of precedence of the scope with respect to the marker and a distribution of the marker and scope across different speech turns and different speakers.

From a functional point of view, it is the creation of a relation between an epistemic marker and a scope that allows the propositional content of the scope to be grounded in the common ground.

Positing epistemic constructions *per se*, rather than epistemic markers, as the object of inquiry has proven to have at least two major advantages. Firstly, it enabled us to apply a unified framework for the analysis of any type of epistemic construction, irrespective of the morphosyntactic nature of the marker or the complexity of the scope. Secondly, it allowed us to take the construction as a domain of analysis and observe the position of epistemic markers within it. Such an approach is quite uncommon since what is usually observed is the position of the marker within a speech turn, a sentence, or an utterance. It has been shown in the

⁴ Consistently with what we proposed in Pietrandrea (2018), we maintain that both epistemic and evidential markers can validate the truth-value of the scope.

⁵ See Pietrandrea (2005, 2018), Boye (2012) for a distinction between propositions as truth-bearer semantic objects and other classes of semantic objects.

literature, though, that these units are not necessarily pertinent to characterize the distribution of an epistemic marker and that the observation of the topological (rather than syntactic) position of a marker with respect to its scope can be more relevant for predicting its function (Masini & Pietrandrea 2010). Hence, by taking the construction, rather than other units, as the domain of analysis, we were able to identify a number of regularities in the association between the position of the epistemic marker and its discourse function, which would have been otherwise overlooked.

3.4 The formal and functional structure of epistemic constructions

Using the data-driven procedure described in Pietrandrea (2018) and Nissim/Pietrandrea (2017), we identified a number of relevant formal and functional properties of the marker, the scope and the relation. As shown in table 1, we specified the illocutionary as well as the morphosyntactic nature of the marker, and the syntactic nature of the scope (whether an utterance or a clause). For the relation, we specified: the linear relation between the marker and the scope (the direction); the distribution of the construction across sources (whether the marker and the scope are uttered by the same speaker or by two different speakers); the type of epistemic meaning encoded by the construction (whether genuinely epistemic or specifying a subtype of evidentiality); the discourse function fulfilled by the construction (whether a qualification or a negotiation of the truth of the scope); the polarity of the scope (positive, negative or neutral).

Marker	Illocution	assertion
		injunction
		question
	Morphosyntax	prosodic profile
		utterance
		list
		pragmatic marker
		ctp
		adverbial
		modal verb
		conditional
		future
		imperfect
Scope	Syntax	utterance
		clause
Epistemic Construction	Direction	m>s
		s>m
		s-m-s
		co-extensive
	Source	same speaker
		other speaker
	Epistemic type	no evidence
		direct-visual
		indirect-inferential
indirect-reportive		
		quotative

	Discourse function	memory	
		qualification	
		Negotiation	acceptation
			non acceptance
			check
			confirmation
			non confirmation
	information		
	Polarity	Positive	
		Negative	
Neutral			

Table 1: The annotation scheme

The theoretical definition, illustration and justification of each of these properties and variables were provided in Pietrandrea (2018) and in Cervoni and Pietrandrea (2018). Consistently with the scope of the special issue, we will focus in the remainder of this article solely on the criteria that we adopted to annotate the morphosyntactic properties of the markers.

4 The epistemic marker

4.1 The epistemic marker as a semantic predicate

We defined above the epistemic marker as a semantic predicate that takes an element of the context (the scope of the epistemic construction) as its argument and validates the truth-value of this argument. In order to explain exactly what a semantic predicate is, we will first briefly describe the framework within which this notion was originally developed, i. e., Meaning Text Theory (Mel'cuk 1988). Meaning Text Theory is a dependency theory, i. e., a theory that recognizes that an asymmetry exists between the words of a sentence, due to the fact that the position, the nature, the very presence of some words – the so-called dependents – are determined by the presence of other words – the heads. For example in a sentence such as (13), the presence, the position and the case of the dependent personal pronouns *he* and *me* are determined by the verb *love*, which has to be regarded as the head of the sentence.

13. He loves me

Building on Tesnière's (1959) original theory of dependency, Meaning Text Theory acknowledges two levels of linguistic dependency: beside a syntactic dependency, it is possible to recognize a level of semantic dependency. Some words, which designate properties or relations, justify, by their presence in the sentence, the occurrence in the same sentence of the words that designate the entities about which those properties or relations are predicated. So for example the verb *loves* which predicates the property of being someone who loves requires, to be saturated, the presence of at least a subject. In this framework the verb *loves* is to be regarded as the semantic predicate, or, in other words, the semantic head, of the sentence (13), while *he* and *me* are its semantic arguments. The semantic dependency is represented as shown in figure 1:

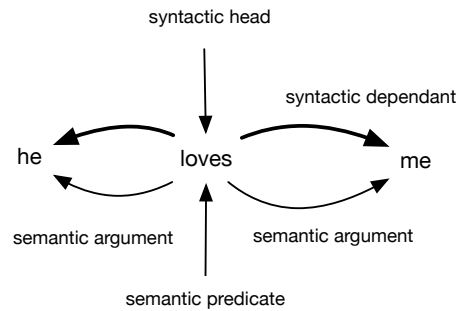


Figure 1

Interestingly enough, Meaning Text Theory recognizes semantic and syntactic dependency as two distinct and autonomous levels of linguistic dependency. As Kahane and Osborne (2015) puts it, in the Meaning Text Theory framework:

semantic dependencies can sometimes point in the same direction as syntactic dependencies and sometimes in the opposite direction (Mel'cuk 1988), e. g. the stone freezes vs. the frozen stone: in both cases the meaning 'freeze' is a predicated of 'stone', so 'stone' semantically depends on 'freeze', whereas stone syntactically depends on freezes, but it governs frozen.

(Kahane and Osborne 2015)

We suggested in Pietrandrea (2018) that epistemic markers are semantic predicates in the sense of Meaning Text Theory, i. e., linguistic elements that designate properties or relations (Polguère 1992; Mel'cuk 2014). As semantic predicates, epistemic markers always govern their scopes semantically (by predicating their truth-values), but from a syntactic point of view, it is entirely possible that sometimes the former may govern the latter whereas sometimes the latter govern the former. Besides, as we will see below, it is also possible for markers and scope to be co-dependent and for them to be syntactically independent from each other. As we will see in the following, it is by taking into account this double level of dependency ((morpho)syntactic and semantic) that we were able to characterize the different types of markers.

4.2 The morphosyntactic nature of epistemic markers

Epistemic markers can be expressed by various morphological, lexical, syntactic or prosodic structures (morphemes, modal verbs, ctp, adverbs, discourse markers, prosodic profiles etc.).

Leaving aside the prosodic profiles, and taking into account the distribution that an epistemic marker has within the dependency tree that represents its syntactic relations with the (syntactic head of its) scope, we distinguished six morphosyntactic classes of epistemic markers: (i) epistemic morphemes, (ii) epistemic modal verb constructions, (iii) CTPs, (iv) epistemic adverbials, (v) epistemic pragmatic markers, (vi) epistemic lists.

- (i) An epistemic morpheme is an affixed element that validates the truth-value of the proposition encoded by the element to which it is affixed. In Italian, epistemic morphemes are always affixed to a verb: see for example, the epistemic future (14), the evidential imperfect (15), and the reportive conditional (16):

14. A: *ma forse non a voi l'**avro**' scritto a un'altra persona*⁶
 'A: but maybe it is not to you **I_might** have written it [lit 3SG.M **have.FUT.1SG** **write.PRTC.PST**] to another person'
 [MoDAL-VOLIP]
15. A: *ma forse non a voi l'**avro**' scritto a un'altra persona*⁷
 'A: but maybe it is not to you **I_might** have written it [lit 3SG.M **have.FUT.1SG** **write.PRTC.PST**] to another person'
 [MoDAL-VOLIP]
16. A: *questo pezzo **sarebbe**?*
 'A: **Is it supposed to be** this piece [lit. this piece **be.COND.PRS.3SG**]?'
 B: *si' tutto questo pezzetto qua*
 'B: yes all this piece here'
 [MoDAL-VOLIP]

As shown in figure 2, epistemic morphemes take the root of the verb expressing the scope as a semantic argument and they depend morphologically on it.

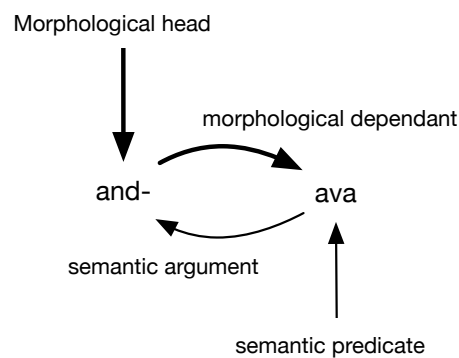


Figure 2

- (ii) An epistemic modal verb construction is a construction characterized by an inflectionally constrained verbal predicate that takes an aspectually incomplete infinitive⁸ as a complement and that validates its truth-value (Pietrandrea 2005; Pietrandrea & Stathi 2010). In Italian modal verb constructions are realized by the indicative (17) and the conditional (18) forms of the verbs *dovere* 'must' and *potere* 'can' followed by incomplete infinitives:

17. *da un momento all'altro io me ne posso anna*
 'At any moment **I can** [lit. **can.PRS.3SG**] pass away'
 [MoDAL-VOLIP]

⁶ For the sake of simplicity, henceforth, we will use a lighter notation of the epistemic constructions: the markers are in bold and the scopes are underlined.

⁷ For the sake of simplicity, henceforth, we will use a lighter notation of the epistemic constructions: the markers are in bold and the scopes are underlined.

⁸ Following Pietrandrea (2005) and Pietrandrea/Stathi (2010) we define incomplete infinitives as infinitives that do not describe changes, or, in more technical terms, as infinitives that can be represented topologically as intervals open to the right (Pietrandrea 2005: 153). Stative, progressive, habitual and resultative infinitives are examples of incomplete infinitives.

18. *in realta' dovrebbero venirci come adesso bollette abbastanza piccole*
 'in fact **we should have** [lit. must.COND.3PL come.INF+2PL.DAT], like now, quite low bills'

[MoDAL-VOLIP]

It has been shown by Pietrandrea (2005) and Pietrandrea and Stathi (2010) that in epistemic modal verb constructions, both the modal verb and the infinitive undergo inflectional constraints. These constructions are therefore better described in terms of co-dependency: they can be regarded indeed as what Gerdes and Kahane (2006) call “verb clusters”, i. e. complex verbal structures consisting of co-dependent verbal elements. The semantic and syntactic dependency structure of epistemic modal verb constructions is represented in Figure 3.

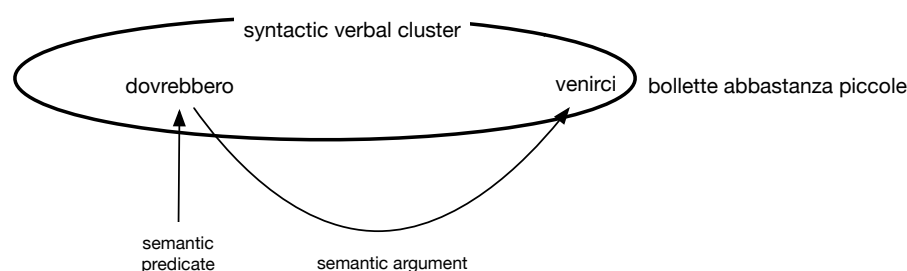


Figure 3

- (iii) A CTP is a verbal predicate that takes a sentential predicate as a complement and that validates its truth-value. Examples of CTPs in Italian are verbs such as *sapere che* ‘to know that’:

19. *noi sappiamo che le donne sono cattive*
 ‘we know that [lit. 1PL.NOM now.PRS.1PL that] women are evil’

[MoDAL-VOLIP]

As shown in figure 4, Epistemic CTPs govern both semantically and syntactically the clausal complement expressing the scope.⁹

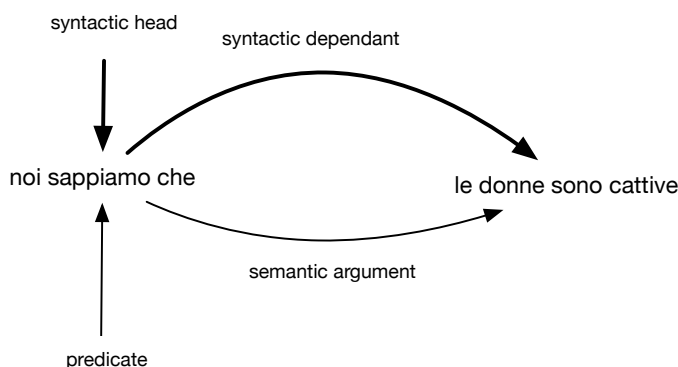


Figure 4

⁹ The distributional criterion that we adopted excludes from the definition the parenthetical uses of CTPs, which, according to our criteria have to be regarded as pragmatic markers, see below.

(iv) An epistemic adverbial is an adjunct (i. e., a syntactically non obligatory) predicate that syntactically depends on another predicate (whether a verb, an adjective, an adverb or an entire clause) and that validates its truth-value. An example of epistemic adverbial is the occurrence of *probabilmente* ‘probably’ in (20):

20. no ma **probabilmente** invece quello che dico io è l’esatto diff<erente> è l’esatto opposto
 ‘no but **probably** what I say is exactly the different is exactly the opposite’

We will see below (section 5.1) what is meant exactly by sentence adverbs, such as epistemic adverbs, being syntactically dependent on their scopes. For the moment let us highlight that epistemic adverbs govern their scopes semantically and are governed by them syntactically.¹⁰ Figure 5 illustrates this state of affairs:

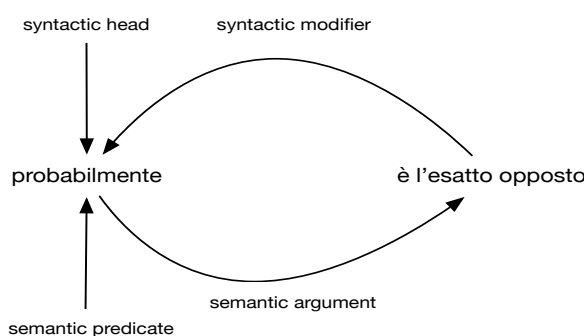


Figure 5

(v) An epistemic pragmatic marker is a predicate that operates on an argument by validating its truth-value and that is syntactically independent of the predicate on which it operates. Examples of epistemic pragmatic markers are *certo* ‘of course’ in (21), *mi sembra* ‘it seems to me’ in (22):

21. *poi io tornavo la sera* *mi sembra*
 ‘then I would come back at night I think [lit. 1SG.DAT seem.PRS.3SG]’
 [MoDAL-VOLIP]
22. A: *ma_ è il periodo del novanta quindi si riferiscono al periodo del novanta?*
 ‘A: but this is 1990, do they refer to 1990?’
 B: **certo**
 ‘B: of course [lit. certain.M.SG]’
 [MoDAL-VOLIP]

As shown in figure 6 the semantic dependency of the scope on the epistemic marker does not correspond to a syntactic dependency in any direction.

¹⁰ When the semantic dependency between two elements points in the opposite direction to the syntactic dependency, one can say that one element modifies the other.

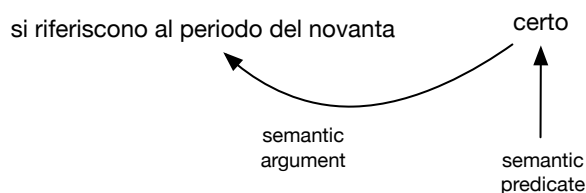


Figure 6

(vi) An epistemic utterance is a syntactically and semantically complete utterance that serves to validate the truth-value of a scope. The reference to the scope is made through a co-reference relation that links an element of the utterance (for example the anaphoric pronoun *lo* ‘that’ in (23)) and (an element of) the scope:

23. *è uno che che ha scritto tra l’altro un articolo* *io non lo sapevo*
 ‘he is someone who wrote an article, by the way’ ‘I didn’t know that [lit. 1SG NEG
 3SG.M know.PST.IMPF.1SG]’

As shown in figure 7 the scope and the marker of a construction realized through an epistemic utterance juxtaposed to an epistemically modalized scope are both syntactically and semantically independent. Their relation is established neither at the syntactic nor at the semantic level, rather it takes place at the discourse level through a co-reference linking:

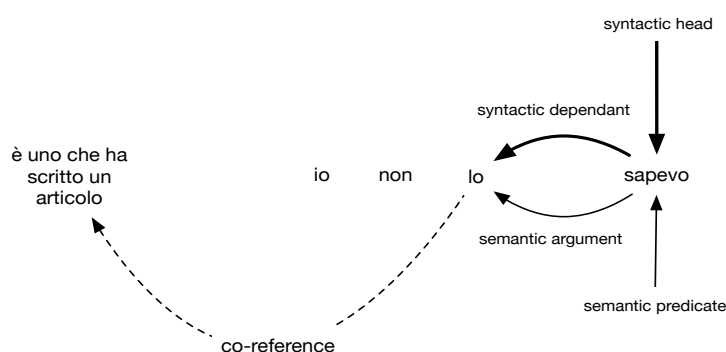


Figure 7

Figure 7 makes clear an important difference between epistemic utterances and epistemic CTPs. The fact that epistemic verbs occur in both epistemic CTPs and epistemic utterances should not make us lose sight of the syntactic differences between the two categories. Epistemic CTPs directly govern their scopes (the scope coincides with the syntactic object of the marker), whereas epistemic utterances do not (the scope is co-referential with the syntactic object of the marker, which is generally realized by an anaphoric pronoun).

(vii) A list construction is a construction “characterized by the syntagmatic co-presence of two or more units potentially standing in a paradigmatic relation with one another, that are not depending on one another and that fill one and the same slot within a syntactic dependency tree” (Blanche-Benveniste 1990; Masini & Pietrandrea 2010; Kahane & Pietrandrea 2012a). An epistemic list construction is a list construction in which the first element of the list represents a scope whose truth-value is validated by further elements of the list.

Two epistemic list constructions can be identified in (24). The first list construction is represented by the two occurrences of the circumstantial adjunct *con i testi* ‘with the texts’: the first occurrence uttered by C represents the scope of the construction, the second occurrence, uttered by A, represents the marker that A uses to check the truth of the scope. The second list is represented by the two occurrences of the circumstantial adjunct *secondo il programma* ‘according to the syllabus’: the first occurrence, uttered by C, represents the scope of the construction, the second occurrence represents the marker that A uses to accept the truth of the scope.

24. C: *l’ho preparato con i testi*
 ‘C: I prepared it with the texts’
 A: **con i testi?**
 ‘A: **with the texts?**’
 C: *sì secondo il programma*
 ‘C: yes, following the syllabus’
 A: **secondo il programma**
 ‘A: **following the syllabus**’

[MoDAL-VOLIP]

As shown in figure 8, in an epistemic list construction the marker governs the scope semantically and it is in a paradigmatic relation with it syntactically.

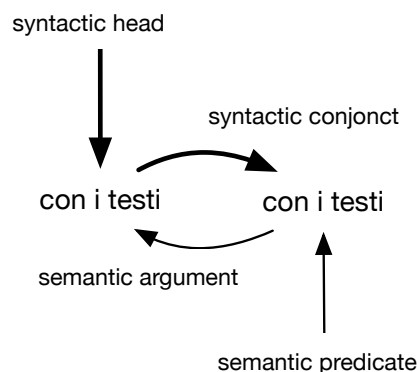


Figure 8

So far we have characterized the morphosyntactic nature of the various types of epistemic markers.

In the next section we will focus on epistemic sentence adverbs and we will characterize them formally, by (i) clarifying what it means for a sentence adverb to be syntactically dependent on its scope; (ii) spelling out the formal difference between epistemic sentence adverbs and epistemic CTPs on the one hand and epistemic sentence adverbs and pragmatic markers on the other hand; (iii) showing what the consequences of a distributional (rather than morphological) classification for the identification of epistemic sentence adverbs in discourse are.

5. Epistemic sentence adverbs

5.1 The syntactic dependency of sentence adverbs

As noted by Quirk et al. (1985) among others, an adverbial can be considered syntactically dependent on another predicate when it can be clefted (25a), interrogated (25b), negated (25c), focused (25d) and coordinated (25e), such as, for example, the adverb *velocemente* ‘quickly’, in (25):

25. *Luigi corre velocemente.*
‘Luigi is running quickly.’
- 25a. *È velocemente che corre Luigi.*
‘It is quickly that Luigi is running.’
- 25b. *Luigi corre velocemente?*
‘Is Luigi running quickly?’
- 25c. *Luigi non corre velocemente.*
‘Luigi is not running quickly.’
- 25d. *Luigi corre proprio velocemente.*
‘Luigi is running really quickly.’
- 25e. *Luigi corre velocemente e maldestramente.*
‘Luigi is running quickly and awkwardly.’

The question arises whether we can consider sentence adverbs as syntactically dependent on their scopes, since sentence adverbs cannot be focalized (Quirk et al. 1985). This entails that they cannot be clefted, they cannot be interrogated, they cannot be negated and they cannot enter the scope of a focalizer; besides, sentence adverbs cannot be coordinated:

26. *Probabilmente Luigi corre.*
‘Probably Luigi is running.’
- 26a. **È probabilmente che corre Luigi.*
‘It is probably that Luigi is running.’
- 26b. **Luigi corre probabilmente?*
‘Is Luigi running probably?’
- 26c. **Luigi non corre probabilmente.*
‘Luigi is not running probably.’
- 26d. **Luigi corre proprio probabilmente.*
‘Luigi is running really probably.’
- 26e. **Luigi corre probabilmente e fortunatamente.*
‘Luigi is running probably and hopefully.’

To put it differently, sentence adverbs are not included in the dependency tree of the sentence that they modify and therefore they cannot be considered *stricto sensu* as syntactically dependent on the head of the sentence. When it was argued above that epistemic adverbs are

syntactically dependent on their scopes and that they are to be distinguished in this respect from epistemic pragmatic markers, this does not mean that they are microsyntactically included in the dependency tree of the sentence. Still, we claim that they are dependent on their scopes at another level of syntactic cohesion, i. e., macrosyntax.

The distinction between micro and macrosyntax was first introduced in the framework of Macrosyntax theory (Blanche-Benveniste 1990; Berrendonner 1990; Cresti 2000) and was recently applied to the analysis of the Rhapsodie corpus (Lacheret et al. 2019; Kahane and Pietrandrea 2019).

In Macrosyntax theory, microsyntax is a cover term to describe two types of cohesion mechanisms defined by the relations between the words of a sequence: government relations and paradigmatic relations. Government relations describe, quite classically, the constraints that a word can operate on a constituent by determining its occurrence, its categorical nature, its position as well as its markers. An example of government relation is the dependency of non-sentence adverbs on the predicates that they modify. Paradigmatic relations are cohesive mechanisms that are orthogonal to government and that guarantee cohesion between the conjuncts of a list (see above), i. e., words occupying one and the same slot in a dependency structure. In other words, microsyntax describes the mechanism of cohesion at work within a dependency tree.

The term macrosyntax, instead, describes the relations between the constituents of an utterance that concur to realize one and only one speech-act. It has been recognized (Blanche-Benveniste 1990; Biber et al. 1999) that within an utterance a distinction can be made between a central, indispensable constituent, which carries the illocutionary force of the utterance and that could be uttered in isolation – the nucleus of the utterance – and a number of optional constituents, which can precede, follow or interrupt the nucleus, which are not endowed per se with an illocutionary force and that cannot therefore be uttered in isolation – the adnuclear constituents. In the answer (27), for example, the nuclear constituent is *l'ho visto in cucina* 'I saw him in the kitchen'; it is endowed with the illocutionary force of an answer and could, as such, be uttered in isolation to answer the question. A number of adnuclear constituents precede *Luigi* (*boh* 'well'), or follow it (*mi sembra* 'I think'); they could not be uttered in isolation and could be removed without affecting the interpretability of the utterance.

27. A: *Dov'è Luigi?* B: *Boh, Luigi, l'ho visto in cucina, mi sembra.*

'A: Where is Luigi? B: Well, Luigi, I saw him in the kitchen, I think'

[MoDAL-VOLIP]

The adnuclear constituents can be regarded as syntactically dependent on the nucleus, since their presence is only allowed by the presence of the nucleus.

With this distinction in mind, it becomes easy to show that epistemic sentence adverbs are macrosyntactically dependent on their scopes, or in other words that they realize adnuclear constituents of the nuclear constituent which realizes their semantic scope. In an utterance such as (20), reproduced here in a slightly simplified form as (28), the illocutionary force of the utterance is carried by the constituent *quello che dico io è l'esatto opposto* 'what I say is

exactly the opposite’, which also realizes the scope of the epistemic construction, while the epistemic sentence adverb *probabilmente* ‘probably’ is an optional adnuclear constituent.

28. *No, ma probabilmente quello che dico io è l'esatto opposto*

‘No, but probably what I say is exactly the opposite’

The asymmetry between the indispensable nuclear constituent *quello che dico io è l'esatto opposto* ‘what I say is exactly the opposite’ and the optional adnuclear constituent *probabilmente* ‘probably’, is shown by the fact that *probabilmente* could not be uttered, in this context, in isolation since its presence is determined by the presence of the nucleus.

We can therefore say that the presence of the adverb depends macrosyntactically on the presence of the predicate it scopes over semantically.

Having characterized the syntactic dependency of sentence adverbs, we can now illustrate the difference between epistemic sentence adverbs and other epistemic markers that partake of some of the properties of sentence adverbs, i. e., epistemic CTPs, epistemic pragmatic markers and epistemic utterances.

5.2 The difference between epistemic sentence adverbs and epistemic CTPs

Epistemic adverbs and epistemic CTPs have sometimes been considered as part of one and the same functional category (Diessel & Tomasello 2001; Thompson 2002).

It should be highlighted, however, that from a syntactic point of view, they show two major differences: firstly, epistemic CTPs are linked microsyntactically rather than macrosyntactically to their scopes; secondly, they govern their scopes rather than being governed by them. Example (19), reproduced here under (29), shows that the scope, and not the marker, can be clefted, interrogated, negated, focalized and coordinated as shown by the manipulations in (29a) through (29e):

29. *Noi sappiamo che le donne sono cattive*

‘we know that [lit. 1PL.NOM know.PRS.1PL that] women are evil’

[MoDAL-VOLIP]

29a. *È che le donne sono cattive quello che noi sappiamo.*

‘It is that women are evil that we know.’

29b. *Noi lo sappiamo che le donne sono cattive?*

‘Do we know that women are evil?’

29c. *Noi non lo sappiamo che le donne sono cattive (sappiamo che sono buone).*

‘We don’t know that women are evil (we know that they are nice).’

29d. *Noi sappiamo proprio che le donne sono cattive.*

‘We know exactly that women are evil.’

29e. *Noi sappiamo che le donne sono cattive e che gli uomini sono buoni.*

‘We know that women are evil and that men are nice.’

This proves that microsyntactic cohesion exists between CTPs and their scopes, which we excluded for sentence adverbs (section 4.1), and that it is the scope that is syntactically dependent on the marker.

5.3 The difference between epistemic sentence adverbs and epistemic pragmatic markers

Let us now examine epistemic pragmatic markers.

30. A: *ma_ è il periodo del novanta quindi si riferiscono al periodo del novanta?*

‘A: but this is 1990, do they refer to 1990?’

B: *certo*

‘B: of course [lit. certain.M.SG]’

[MoDAL-VOLIP]

As example (22), reproduced here as (30), shows, unlike *probabilmente* in (30), the marker *certo* realizes, syntactically speaking, an autonomous constituent. From a microsyntactic point of view, it does not show any dependency on its nucleus (it would not meet any of the criteria of dependency illustrated above in 25a through 25e). From a macrosyntactic point of view, it is equally clear that *certo* is endowed with its own illocutionary force (it realizes an answer) and that it can, in fact it is, uttered in isolation. The relation between the marker *certo* and its scope *si riferiscono al novanta*, ‘they refer to 1990’ is therefore purely semantic: *certo* is a semantic predicate whose semantic argument is realized by the *si riferiscono al novanta*, ‘they refer to 1990’. From a syntactic point of view *certo* and its scope are completely independent.

It should be highlighted that beside interjections like *certo*, we include the so-called parenthetical uses of CTPs in the category of epistemic pragmatic markers (31):

31. A: *l'Anita ha deciso di stare con tutt'e due* D: *ho capito*

‘A: Anita decided to date both of them D: I see’

[MoDAL-VOLIP]

As shown by Kahane and Pietrandrea (2012b), these elements, which are by the way often included in the category of pragmatic markers for functional reasons (see Schiffrin 1987, among others), also show important syntactic similarities with elements such as *certo* examined above. Parenthetical verbs, in fact, scope over their arguments semantically and they are both micro- and macro-syntactically independent of them: they would not pass the tests of syntactic integration illustrated in (25a) through (25e). and, since they are endowed with their own illocutionary force, they can be uttered in isolation as shown in (33), and are therefore also macrosyntactically autonomous.

5.4 A distributional classification

It is important to emphasize that the classification that we have proposed above is not a classification of forms; rather, it is a classification of distributions. This means that it is entirely possible, as the examples below show, for the same word to have different distributions and therefore for it to realize different syntactic functions. For instance, *forse* ‘maybe’ in (32) is employed as an adverb in (32a) and as a pragmatic marker in (32b); *capire*

‘to understand’ is used as a CTP in (33a) as a pragmatic marker in (33b) and in an autonomous epistemic utterance in (33c):

32a. *forse lei potrebbe fare meglio*
‘maybe you could do better’

[MoDAL-VOLIP]

32b. A: *È uscito?* B: *forse*
‘A: Has he gone out? B: maybe’

[MoDAL-VOLIP]

33a. *ho capito che è un caso che ha appassionato molto*
‘I understood that this was a thrilling case’

[MoDAL-VOLIP]

33b. A: *l’Anita ha deciso di stare con tutt’e due* D: *ho capito*
‘A: Anita decided to date both of them D: I understood (roughly I see)’

[MoDAL-VOLIP]

33c. D: *cos’è che non ti ricordi?* B: *un cazzo* D: *ah non l’avevo capito*
‘D: what is it that you don’t remember? B: nothing D: Ah I had not understood that’

[MoDAL-VOLIP]

It should be highlighted, though, that some words seem to be uniquely associated with a given function. For example, the data in the ItTenTen Corpus – a corpus made up of texts collected from the Internet, which is part of the TenTen family and which includes 4.9 billion words (Jakubíček et al. 2013) – showed that the markers *secondo me*, *in realtà*, *a dire la verità*, never occur in positions other than the adverbial one.

6. Functional differences between epistemic sentence adverbs and other epistemic markers

So far we have distinguished epistemic adverbs from epistemic pragmatic markers and CTPs on a purely syntactic basis. The question arises whether this formal classification can also be functionally meaningful: Do the different morphosyntactic classes fulfill different functions? And if so, at what level?

6.1 A general picture

The quantitative analysis of our corpus has provided some answers. Before presenting the main findings, let us offer a general picture of the quantitative distribution of the various types of epistemic markers in the corpus.

Morphosyntactic types	Occurrences	Proportion
Epistemic pragmatic markers	324	45%
Epistemic CTPs	162	22%
Epistemic lists	82	11%
Epistemic prosodic profiles	48	7%
Epistemic utterances	46	6%
Epistemic adverbs	37	5%

Epistemic modal verbs	12	2%
Epistemic morphemes	11	2%

Table 2: Distribution of the morphosyntactic types of epistemic markers

As shown in table 2, epistemic pragmatic markers and epistemic CTPs are by far the most frequent morphosyntactic markers of epistemicity (together they represent 67% of the occurrences of epistemic markers); likewise, epistemic adverbs, albeit less frequent (5% of the occurrences), are still much more frequent than epistemic modal verbs or epistemic morphemes, which are the most well known and most widely studied markers of epistemicity. This means two things: (i) we have enough data to put forward hypotheses that are worthwhile exploring and, more importantly, (ii) we are not talking about a marginal phenomenon, rather we are dealing with a class of markers which is quite central for the analysis of epistemicity.

6.2 Functional properties of epistemic sentence adverbs

6.2.1 Functional differences between epistemic sentence adverbs and epistemic CTPs

It has already been shown in the literature that there is a functional distinction between CTPs and epistemic sentence adverbs. Boye and Harder (2007) showed for example that CTPs do not simply behave as clause operators, rather they can potentially function both as foregrounded addressable descriptions of the epistemic evaluation (34) and as expressions of an evaluation put forward as backgrounded and non addressable in discourse (35).

34. *Si ma noi sappiamo che le donne sono cattive lo sappiamo in un punto molto preciso storico*

‘Yes but we know that women are evil. We learned that at a precise point, a historical moment’

[MoDAL-VOLIP]

35. *E: ma che mori me sa che moro prima io che te moro ma come fa*

‘E: Why do you say that you will die I think I will die before you die, but how can you?’

[MoDAL-VOLIP]

D: ah perche' io ho fatto il contratto col padreterno insomma

‘D: ah, so you think that I have a contract with the Eternal Father, don’t you?’

[MoDAL-VOLIP]

As hypothesized by Boye and Harder and confirmed by our corpus data, CTPs are used as flexible evaluation devices. The fact that they may function both as descriptions of foregrounded evaluations and as expressions of backgrounded evaluations provides a resource for the speaker who can present a complex sentence comprising both an evaluating and an evaluated part, allowing the hearer to address either the former (as in (34)) or the latter (as in (35)) if necessary. This functional property of CTPs is not shared by adverbs, which are invariably presented as backgrounded evaluations and are therefore never addressed in discourse.

6.3 Functional differences between epistemic sentence adverbs and epistemic pragmatic markers

As far as the distinction between epistemic sentence adverbs and epistemic pragmatic markers is concerned, the analysis of the Modal corpus seems to suggest that such a distinction is pertinent on the functional level as well. Namely, the two classes of epistemic markers seem to provide different contributions to dialogue construction.

As Table 3 and 4 show, 86.5% of epistemic adverbials and 84.5% of epistemic CTPs are used to qualify rather than negotiate the truth-evaluation of their scopes, whereas 88% of pragmatic markers are used to negotiate rather than qualify the truth-evaluation of their scopes. This distribution is significant at $p < 0.00001$.

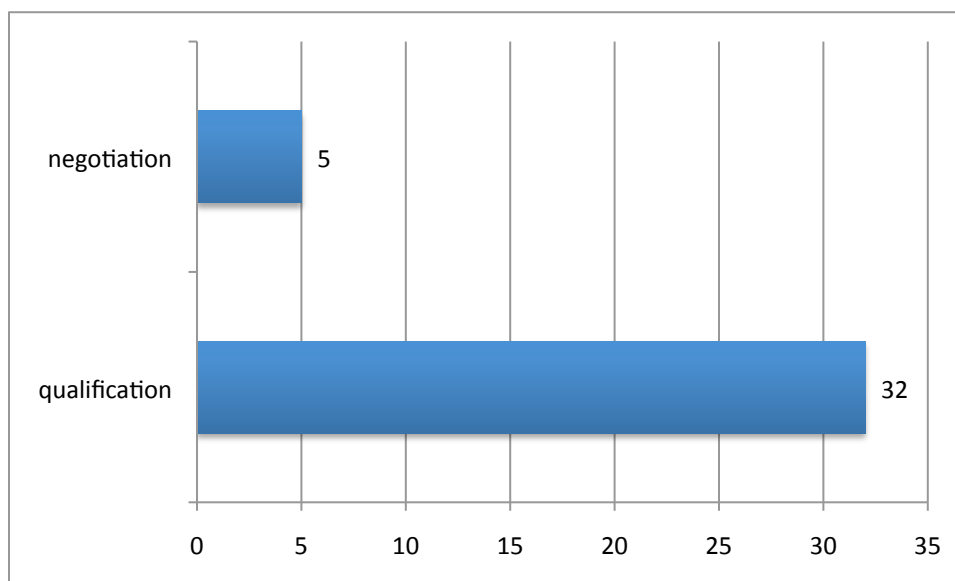


Table 3: The discursive functions of adverbs

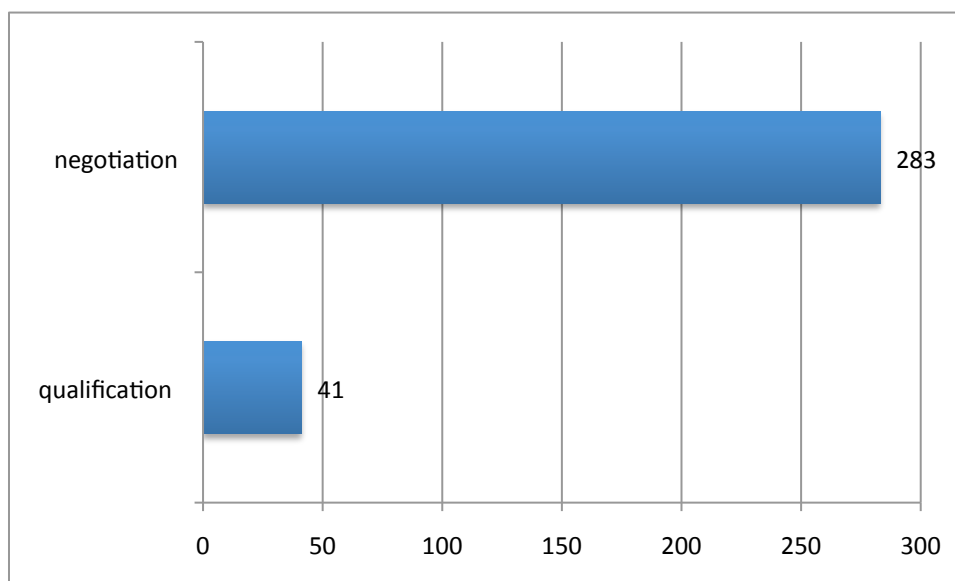


Table 4: The discursive functions of pragmatic markers

The opposition between epistemic qualification and epistemic negotiation is represented in the annotation scheme of the Modal corpus and thoroughly justified in Pietrandrea (2018). In

this framework, an epistemic construction is considered as epistemically qualifying its scope when the epistemic marker is merely used to indicate the knowledge or evidence the speaker has to validate the truth of a scope, which he wants to add to the common ground.

36. *io non credo che io abbia fatto un contratto*
 ‘I don’t think that [lit. 1SG.NOM believe.PRS.1SG that] I have a contract’
col Padreterno
 ‘with the Eternal Father’
 [MoDAL-VOLIP]

An epistemic construction is considered as epistemically negotiating its scope, instead, when the marker allows for a dialogical negotiation of the validation of the truth-value of a commitment, which includes acceptance, non-acceptance, check, confirmation, non confirmation, and information.

37. B: *sentia io pure ci ho ci ho ma<mma> mia madre che e’ morta nove mesi fa* A: *ah*
 ‘B: listen, but my mum my mother died nine months ago A: oh’
 [MoDAL-VOLIP]
38. A: *mi sembra all’hotel Cento Stelle bah?* B: *mh*
 ‘A: I think at the hotel Cento Stelle, right? B: maybe [lit. mh]’
 [MoDAL-VOLIP]
39. *la cosa nuova e’ un disastro capito?*
 ‘the new thing is a disaster you know [lit. understand.PTCP.PST.M]?’
 [MoDAL-VOLIP]
40. A: *e’ un problema? ha un problema?* B: *si*
 ‘A: is it a problem? do you have a problem? B yes’
 [MoDAL-VOLIP]
41. A: *avete fissato i termini?* B: *no i termini no*
 ‘A: did you set the deadline? B: No not the deadline’
 [MoDAL-VOLIP]

7 Conclusions

We have shown that, at least as far as the epistemic domain is concerned, it is possible to identify an autonomus class of epistemic sentence adverbs. Epistemic sentence adverbs can be regarded as semantic predicates that validate the truth value of their scopes and that are macrosyntactically dependent on them.

In this respect, epistemic sentence adverbs should be distinguished on the one hand from epistemic CTPs, which are defined as semantic predicates that validate the truth value of their scopes and govern them microsyntactically, and on the other hand from epistemic pragmatic markers, which are defined as semantic predicates that validate the truth value of their scopes and that are syntactically independent of them.

Such a distinction is not necessarily inscribed at the lexical level, so, even though we identified some words that invariably behave as sentence adverbs, other words can have

various distributions and can therefore be classified as epistemic sentence adverbs in some syntactic contexts, and as other epistemic markers in other syntactic contexts.

We have hypothesized that epistemic sentence adverbs also have a specific functional role. Unlike (most) epistemic CTPs they express a non-addressable epistemic evaluation of the scope; unlike pragmatic markers they serve to qualify, rather than to negotiate, the epistemic evaluation of the scope. In order to confirm this hypothesis a thorough analysis of a corpus annotated for epistemic constructions, information status and dialogical moves is currently being performed.

Corpora

VoLIP Corpus. <http://www.parlaritaliano.it/index.php/it/VoLIP>

Santa Barbara Corpus. <http://www.linguistics.ucsb.edu/research/santa-barbara-corpus>

ESLO Corpus. <http://eslo.huma-num.fr>

Accueil UBS Corpus. http://www.info.univ-tours.fr/~antoine/parole_publicue/Accueil_UBS/index.html

MoDAL Corpus. <http://modal.msh-vdl.fr/?lang=en>

OTG Corpus. http://www.info.univ-tours.fr/~antoine/parole_publicue/OTG/index.html

References

- Berrendonner, Alain (1990): « Pour une macro-syntaxe ». *Travaux de linguistique* 21: 25–36.
- Biber, Douglas/Johansson, Stig/Leech, Geoffrey/Conrad, Susan/Finegan, Edward (1999): *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education Limited.
- Blanche-Benveniste, Claire (1990): « Un modèle d'analyse syntaxique 'en grilles' pour les productions orales ». *Anuario de Psicología* 47: 11–28.
- Boye, Kasper (2012): *Epistemic Meaning. A Cross-Linguistic and Functional Cognitive Study*. Berlin: Mouton de Gruyter.
- Boye, Kasper/ Harder, Peter (2007): “Complement taking predicates : usage and linguistics structures”. *Studies in Language* 31(3): 569–606.
- Cervoni, Valerio/Pietrandrea Paola (2018): « Equivalenze pragmatiche tra i marcatori epistemici del francese e dell'italiano parlato ». . In: A. De Meo, F. Dovetto SLI – GSCP International Conference Naples 13–15 June 2016. 209-229
- Cresti, Emanuela (2000): *Corpus di italiano parlato*. Florence: Accademia della Crusca.
- Diessel, Holger/Tomasello, Michael (2001) : “The acquisition of finite complement clauses in English: A corpus-based analysis”. *Cognitive Linguistics* 12: 1–45.
- Gerdes Kim/Kahane, Sylvain (2006): « L'amas verbal au cœur d'une modélisation topologique de l'ordre des mots ». *Linguisticae Investigationes* 29 (1): 75–89.
- Jakubíček, Milos/Kilgarriff, Adam/Kovář, Vojtěch /Rychlý, Pavel/Suchomel, Vit (2013): “The TenTen corpus family”. In: Proceedings of the 7th International Corpus Linguistics Conference. Lancaster, July: 125–127.
- Kahane, Sylvain/Osborne, Otto (2015): “Translators' introduction”. Lucien Tesnière (1959) *Elements of Structural Syntax*. Amsterdam : Benjamins.
- Kahane, Sylvain/Pietrandrea, Paola (2012a): « La typologie des entassements en français ». In: Actes du 3e congrès mondial de linguistique française (CMLF). Lyon, July 2012: 1809–1828.

- Kahane, Sylvain/Pietrandrea, Paola (2012b): « Les parenthétiques comme Unités Illocutoires Associées. Une perspective macrosyntaxique ». *Linx* 61 (2009): 49–70.
- Kahane, Sylvain/Pietrandrea, Paola (2019): “The syntactic annotation of the Rhapsodie corpus: an overview”. In: Lacheret, Anne/Kahane, Sylvain/Pietrandrea, Paola (eds.): *Rhapsodie: a prosodic syntactic treebank of spoken French*. Amsterdam/Philadelphia: Benjamins.
- Kärkkäinen, Elise (2012): “I thought it was very interesting. Conversational formats for taking a stance”. *Journal of Pragmatics* 44 (15): 2194–2210.
- Lacheret, Anne/Kahane, Sylvain/Pietrandrea, Paola (eds.) (2019): *Rhapsodie: a prosodic syntactic treebank of spoken French*. Amsterdam/Philadelphia: Benjamins.
- Masini, Francesca/Pietrandrea, Paola (2010): “Magari”. *Cognitive Linguistics* 21 (1): 75–121.
- Mel’čuk, Igor (1988): *Dependency syntax: Theory and practice*. Albany, NY: SUNY Press.
- Mel’cuk, Igor (2014): “Dependency in Language”. In: Wright, James (ed.), *International Encyclopedia of the Social & Behavioral Sciences*. 2nd Edition. Vol. 6. Oxford, Elsevier: 182–195.
- Mithun Marianne (2012): “Tags: Cross-linguistic diversity and commonality”. *Journal of Pragmatics* 44 (15): 2165–2182.
- Nissim, Malvina/ Pietrandrea, Paola (2017): “MODAL: A multilingual corpus annotated for modality”. In: Proceedings of the 4th Italian Conference on Computational Linguistics (CLiC-it 2017). Rome, Italy, December 11-13, 2017.
- Pietrandrea, Paola (2005): *Epistemic Modality. Functional Properties and the Italian System*. Amsterdam/Philadelphia: Benjamins.
- Pietrandrea, Paola (2018): “Epistemicity at work. A corpus study on Italian dialogues”. *Journal of Pragmatics* 128 : 171-191.
- Pietrandrea, Paola/Stathi, Katerina (2010): “What counts as an evidential unit? The case of evidential complex constructions in Italian and Modern Greek”. *STUF - Sprachtypologie und Universalienforschung* 63 (3): 333–344.
- Polguère, Alain (1992): « Remarques sur les réseaux sémantiques Sens-Texte ». In: Clas, André (ed.): *Le mot, les mots, les bons mots*. Montreal, Les Presses de l’Université de Montréal: 109–148.
- Quirk, Randolph/Greenbaum, Sidney/Leech, Geoffrey/Svartvik, Jan (1985): *A Comprehensive Grammar of the English Language*. London: Longman.
- Rossari, Corinne (2012): “The evidential meaning of modal parentheticals”. *Journal of Pragmatics* 44 (15): 2183–2193.
- Schiffrin, Deborah (1987): *Discourse Markers*. Cambridge: Cambridge University Press.
- Simon-Vandenberghe, Anne Marie/Aijmer, Karin (2007): *The semantic field of modal certainty: a corpus-based study of English adverbs*. Berlin/New York: Mouton de Gruyter.
- Tesnière, Lucien (1959): *Éléments de syntaxe structurale*. Paris: Klincksieck.
- Thompson, Sandra A. (2002): “Object complements and conversation: towards a realistic account”. *Studies in Language* 26 (1): 125–164.
- Traugott Elizabeth (2012): “Intersubjectification and clause-periphery”. *English Text Construction* 5 (1): 7–28.