



Designing and validating an annotation model for dynamic modality in the MULTINOT project: issues and problems

Julia Lavid, Marta Carretero and Juan Rafael Zamorano-Mansilla

Universidad Complutense de Madrid

julavid@filol.ucm.es, mcarrete@filol.ucm.es,

juanrafaelzm@filol.ucm.es

Abstract

In this paper we set forth an annotation model for dynamic modality in English and Spanish, given its relevance not only for contrastive linguistic purposes, but also for its impact on practical annotation tasks in the Natural Language Processing (NLP) community. An annotation scheme is proposed, which captures both the functional-semantic meanings and the language-specific realisations of dynamic meanings in both languages. The scheme is validated through a reliability study performed on a randomly selected set of one hundred and twenty sentences from the MULTINOT corpus, resulting in a high degree of inter-annotator agreement. We discuss our main findings and give attention to the difficult cases as they are currently being used to develop detailed guidelines for the large-scale annotation of dynamic modality in English and Spanish.

1 Introduction

This paper reports on current work on the annotation of modality meanings in English and Spanish in the context of the MULTINOT project, aimed at the creation of a high-quality, register-diversified parallel and medium-sized corpus for the English-Spanish pair. The MULTINOT corpus consists of originals and translated texts in both directions and is enriched with linguistic annotations which can be exploited in a number of linguistic, applied and computational contexts (see Lavid et al. 2015).¹ More specifically, in this paper we focus on one of the subtypes of modality studied in the literature, i.e., the so-called *dynamic modality*, given its relevance not only for theoretical and descriptive purposes, but also for its impact on practical annotation tasks in the Natural Language

¹ The MULTINOT project is financed by the Spanish Ministry of Economy and Competitiveness, under grant number FF2012-32201. The authors of this paper gratefully acknowledge the support provided by the Spanish authorities.

Processing (NLP) community. In fact, applications such as textual entailment, information extraction, question answering, sentiment analysis and machine translation need to be able to distinguish automatically modal from actual information as part of the complete understanding of a text. Thus, researchers in the NLP community have developed annotation schemes and annotated corpora for different aspects of modality in different languages in the last years (see McShane et al., 2004; Wiebe et al., 2005; Szarvas et al., 2008; Sauri and Pustejovsky, 2009; Hendrickx et al., 2012; Baker et al., 2012, among others), which can be used as training data for the development of computational systems which automatically distinguish modal information in texts.

However, the annotation of modal meanings is not an easy task, and the area of dynamic modality poses specific problems for practical annotation, especially when dealing with different languages such as English and Spanish, as will be shown in the rest of this paper. Moreover, there are not annotated datasets available for these two languages in terms of modality meanings, so with our work we hope to contribute to current research efforts in this area, by providing an annotation model for dynamic modal meanings in English and Spanish which will be the basis for the larger-scale annotation of the different registers included in the bilingual MULTINOT corpus.

The paper is organised as follows: in section 2 we review the notion of dynamic modality in the literature and discuss the main problems and issues surrounding this notion. Section 3 presents our own proposal for capturing dynamic meanings in English and Spanish in the form of an annotation scheme consisting of two tagsets: a) a functional/semantic tagset, capturing the functional/semantic meanings of dynamic modality, and, b) a lexicogrammatical tagset, capturing the syntagmatic options realising those meanings in English and Spanish. The functional tagset is subject to empirical validation through a pilot reliability study performed on a sample consisting of sentences extracted from the MULTINOT corpus. In section 4 we discuss the problematic cases and the difficulties encountered as the result of the pilot study. Finally Section 5 summarises the work being reported and provides some pointers for future work.

2 Dynamic modality: problems and considerations

The label ‘dynamic modality’ is widely used in the literature to refer to a semantic subtype of modality (Hermerén, 1978; Palmer, 1990; Perkins, 1983; Carretero, 1995; Silva-Corvalán, 1995; Nuyts, 2001, 2005; Wårnsby, 2006; Collins, 2009; Portner, 2009; Loureiro-Porto, 2013). This term, together with ‘epistemic modality’, ‘deontic modality’ and less commonly ‘boulomaic/volitional modality’, ‘evaluative modality’ and others, corresponds to a typology of modal meanings based on modal logic and centered around the notions of possibility and necessity. The references cited above display minor differences about the concept and scope of dynamic modality, but basically coincide in their description of this category as covering the notions of *physical necessity*, *habit* or *tendency*, *ability* and *non-inherent (or extrinsic) possibility*. We follow the proposal set forth in Perkins’s (1983) monograph on modal expressions in English, due to its clarity for distinguishing between types of modality. Perkins’s approach was also adopted in Silva-Corvalán’s (1995) study of the Spanish periphrases with *poder* (‘can, may’) and *deber (de)* (‘must’), and by earlier work by two of the authors of this paper (Zamorano-Mansilla and Carretero 2012), who address the annotation of a limited number of English and Spanish polysemous modal expressions, namely the English modal auxiliaries *can*, *must* and *have to*, their Spanish equivalents *poder*, *deber (de)* and *tener que* (‘have to’), the English adverb *possibly* and its Spanish equivalent *posiblemente*.

Perkins’s approach is based on Rescher’s (1968) system of modal logic, which consisted of eight categories (alethic, epistemic, temporal, boulomaic, deontic, evaluative, causal and likelihood), which he reduces to three: epistemic, deontic and dynamic. Perkins describes the meaning of modal expressions according to the following elements:

- a) A system of organized beliefs K, which is divided into rational laws (for epistemic modality), social laws (for deontic modality) and natural laws (for dynamic modality).
- b) A set of circumstances C, in which K is relevant.
- c) A variable X, which represents the truth of a proposition in the case of epistemic modality and the occurrence of an event in the case of deontic and dynamic modality.
- d) An axis of three degrees of strength, ordered as follows from the strongest to the weakest: K (C does not preclude X); K (C is disposed towards X); and K (C entails X).²

That is to say, Perkins describes dynamic modality in terms of natural laws, while epistemic modality covers rational laws (in other words, degrees of probability) and deontic modality pertains to social laws (obligation and permission). To this distinction we must add, however, that natural laws do not only cover nature proper, but also states of knowledge and possibilities resulting from the development of science by man. The distinction between natural laws and social laws is made clear by the fact that only social laws can be infringed. Accordingly, *can't* is dynamic in 'Humans can't run as fast as lynces', and deontic in 'We can't smoke in public buildings in Spain'.

An example of strong dynamic modality is (1), since *inevitable* indicates that natural laws (our present state of knowledge) bind us humans to inaccuracy in doing maps of distant areas of the Cosmos. A case of medium dynamic modality is (2), where *tend* indicates that natural laws (the laws of nature proper) are biased towards the survival of the good genes. In its turn, (3) is an instance of weak dynamic modality, where *can* indicates that natural laws (again our state of knowledge) do not preclude us to detect metals in the way described in the clause:

- (1) In Eratosthenes' time, globes were constructed portraying the Earth as viewed from space; they were essentially correct in the well-explored Mediterranean but became more and more inaccurate the farther they strayed from home. Our present knowledge of the Cosmos shares this disagreeable but **inevitable** feature. (Multinot: EO_EXPE_002)
- (2) Because the world has a certain stability and doesn't change capriciously, the genes that have survived in the past **tend** to be the ones that are going to be good at surviving in the future. (Multinot: EO_EXPE_006)
- (3) Today metals **can** be detected in all manner of matrices, down to parts per billion levels – but with a wide range of techniques available, it can be hard to choose the right instrument for the job. (Multinot: EO_POPSCI_003)

Within dynamic modality, we distinguish the meanings of *necessity*, *tendency* and *possibility*, which correspond to the strong, medium and weak degrees, respectively. Within *possibility*, we have considered it convenient to distinguish the subcategory of *ability*, in accordance with many references (Palmer 1990, Portner 2009). *Ability* concerns a skill that someone has acquired (normally in a voluntary way) or that has been conferred to an inanimate entity, as in (4). The other meaning of dynamic weak modality, which concerns the possibility for something to occur due to extrinsic factors (circumstances) rather than participants, has been labelled *situational possibility* (5):

- (4) There is also a 'converse' to this in a system where the computer simulates a human schizophrenic patient, giving all the textbook answers and symptoms, and is **capable** of fooling some medical students into believing that a human patient is actually supplying the answers! (Multinot: EO_EXPE_004)

² Perkins (1983) also proposes a difference in meaning between the primary and the secondary modal auxiliaries, but we do not describe it here for reasons of space.

- (5) Tune your television to any channel it doesn't receive, and about 1 percent of the dancing static you see is accounted for by this ancient remnant of the Big Bang. The next time you complain that there is nothing on, remember that you **can** always watch the birth of the universe. (Multinot: EO_EXPE_001)

The category of dynamic modality poses problems for research, for several reasons. One of them is its peripheral status in the literature within the area of modality. As shown in Zamorano-Mansilla and Carretero (2012: 304-305), some scholars do not establish a distinction between deontic and dynamic modality, and simply use the term 'root modality' for both. In their turn, Halliday and Matthiessen (2014: 696) state that the category of ability / potentiality is 'on the fringe of the modality system'. This frequent consideration of dynamic modality as a peripheral category in the literature is due, in all probability, to its lower degree of subjectivity in comparison to epistemic and deontic modality, both of which are easily perceived as expressing attitudes of the speaker or writer towards what is communicated. Dynamic modality can also be considered as subjective in the sense that its meanings are not directly perceivable; however, these meanings can easily be inferred by the perception of states or events and therefore leave less room for interpersonal variation. For example, there is likely to be much more agreement in the assessment of the number of languages that a person is able to speak (dynamic modality) than in the assessment of his/her probability of winning a literature prize (epistemic modality).

Another problem of dynamic modality is posed by the occurrences in which it communicates conversational implicatures close to the meanings of other modalities. A typical case is 'Can you pass me the salt?' (Groefsema, 1992), which literally expresses *situational possibility* but implicates a request to carry out the action, with the consequent association to deontic modality. This problem occurs above all in spoken language and its recreation in written language (for example, dialogues in fiction and drama).

Finally, in many cases it is problematic to distinguish between dynamic modality and other modal categories, and also between the subtypes of dynamic modality specified above. Some of these problems were encountered in our reliability study and are reported in Section 4.

3 Annotation proposal

Once the conceptual domain of dynamic modality and its subcategories have been described, in this section we present the annotation scheme that we have designed for English and Spanish, which consists of two interrelated tagsets: a) a core functional/semantic tagset, capturing the basic meaning distinctions included in dynamic modality, and, b) a lexicogrammatical tagset, capturing the syntagmatic options realising those meanings in English and Spanish. Next, we describe the reliability study which we have carried out in the context of the MULTINOT project in order to validate the tagsets proposed.

3.1 Annotation tagsets

In line with previous work in the area of epistemicity in English and Spanish (see Lavid et al. 2016), we propose to use two interrelated tagsets to be able to identify the functional similarities and the linguistic differences between these two languages in the area of dynamic modality. On the one hand, we propose a functional/semantic tagset, which captures the dynamic meanings that occur both in English and in Spanish, as graphically displayed in table 1 below:

	Necessity [NE]	
--	----------------	--

Dynamic [DY]	Tendency [TE]	
	Possibility [PO]	Ability [PO_AB]
		Situational Possibility [PO_SI]

Table 1: Functional/semantic Tagset for dynamic modality in English and Spanish

As shown above, the tagset is hierarchical, allowing annotators to choose more general or coarser tags when in doubt about the more delicate ones. For example, if the annotator is uncertain about whether a markable is ‘*situational possibility*’ or ‘*ability*’, s/he can simply tag it as ‘*possibility*’. The abbreviated form of each tag is given in capital letters in brackets next to the full form.

On the other hand, we propose a linguistic tagset, which captures the language-specific realisations of the dynamic meanings presented in table 1 above. The tags here capture a wide variety of linguistic realisations in English and Spanish both in terms of lexicogrammatical options (LG) and in terms of the syntactic functions and constructions (SF) where the lexicogrammatical options can occur, as shown in table 2:

LEXICO-GRAMMAR	SYNTACTIC FUNCTION /CONSTRUCTION	ENGLISH	SPANISH
Adverb [A]	Modal Adjunct [AD]	<i>Inevitably, unavoidably, inexorably, possibly,</i>	<i>Inevitablemente, inexorablemente,</i>
[AJ]	Predicative Adjective in impersonal matrix clause [AJIP]	<i>It is unavoidable/inevitable/bound to happen + that</i>	<i>Es inevitable, ineludible, inexorable que + Subjunctive</i>
	Predicative Adjective in to+ infinitive clause [AJIF]	<i>He is liable, apt to become angry over small things.</i>	<i>Él es propenso, proclive, tendente a...</i>
	Predicative Adjective in impersonal to+infinitive clause	<i>It is inevitable/unavoidable + to + Infinitive</i>	<i>Es inevitable ponerse nervioso en público</i>
	Attributive Adjective in NG [AJN]	<i>He is a potential winner</i>	<i>Es una apuesta segura/</i>
[N]	Noun complement in impersonal matrix clause [NI]	<i>There is a tendency/liability that + Indicative</i>	<i>Hay (una) tendencia/predisposición, a que + Subjunctive</i>
[V]	Verbal operator in matrix clause [VO]	<i>Oil will float on water. This paint can be applied with a spray.</i>	<i>Esa pintura puede aplicarse con un spray.</i>
	Verbal inflection [VI]	----	<i>Los chicos siempre serán chicos</i>

Table 2: Linguistic Tagset for dynamic modality in English and Spanish

The lexicogrammatical options are specified as a core tagset capturing the paradigmatic and more general linguistic encodings of dynamic meanings in English and Spanish (i.e., as adverb, adjective, noun or verb). The syntactic functions and constructions are specified as an extended tagset capturing the syntagmatic encodings where the lexicogrammatical options can occur in both languages. Some tags only hold for one of the languages (i.e., verbal inflection only holds for Spanish). In such cases, we provide an example of the available language and cross out the one that is not available in the other language.

3.2 Reliability Study

In order to test the reliability and consistency of the functional tagset proposed for annotating dynamic meanings in English and Spanish, we carried out a reliability study on a randomly selected set of one hundred and twenty sentences from the MULTINOT corpus (seventy sentences in English and fifty in Spanish). The sentences contained lexicogrammatical candidates which can typically express the dynamic meanings mentioned in Section 2, such as *necessity* (6), *tendency* (7), *possibility* (8), or not express modality at all. For example, ‘liability’ may express tendency, as in ‘His medical history shows a liability to thromboses’, or not express dynamic modality at all, as in (9), where ‘liabilities’ means ‘debts’:

- (6) None of this is easy. These are challenging times, a real stress test for the EU. The path of permanent and profound reform is as demanding as it is **unavoidable**. Let's make no mistake: there is no way back to business as usual. (Multinot: EO_SPEECH_001)
- (7) Yet in Europe we are **prone** to look inwards. We are too defensive about changes in the world. Most of us expect our economic prospects to deteriorate in the years ahead. (Multinot: EO_SPEECH_003)
- (8) Consider these facts: by the most cautious estimates, 400 million people lack the calories, protein, vitamins and minerals needed to sustain their bodies and minds in a healthy state. Millions are constantly hungry; others suffer from deficiency diseases and from infections they would be **able** to resist on a better diet. (Multinot: EO_ESSAY_010)
- (9) Gradual fiscal consolidation – reducing the projected future size of government spending, and hence future tax rates – will have to be at the center of the effort. This should be combined with the mutualization of some portion of the **liabilities** of highly indebted countries. (Multinot: EO_ESSAY_005)

The lexicogrammatical candidates included equal proportions of adjectives, nouns, adjectives, lexical verbs and modal verbs. The annotations were carried out by two expert annotators who tagged both the English and the Spanish sentences independently. Inter-annotator agreement results for the Spanish sentences are presented in table 3:

		Annotator A					
		NON-MODAL	NECESSITY	TENDENCY	SITUATIONAL POSSIBILITY	ABILITY	MODAL (OTHER)
Annotator B	NON-MODAL	0	1	0	0	0	0
	NECESSITY	0	12	0	0	0	0
	TENDENCY	0	0	9	0	0	0
	SITUATIONAL POSSIBILITY	0	0	0	11	3	0
	ABILITY	0	0	0	2	9	0
	MODAL (OTHER)	0	2	0	1	0	0

Table 3: Inter-annotator agreement for Spanish examples

The number of observed agreements for Spanish examples is 41 (82% of the observations), and the number of agreements expected by chance is 11.8 (23.56% of the observations). The kappa value is 0.765. Therefore, the strength of the agreement is considered to be high. In the case of the English sentences, the inter-annotator agreement results are presented in table 4 below:

		Annotator A					
		NON-MODAL	NECESSITY	TENDENCY	SITUATIONAL POSSIBILITY	ABILITY	MODAL (OTHER)
Annotator B	NON-MODAL	0	1	0	2	0	0
	NECESSITY	0	6	0	0	0	0
	TENDENCY	0	0	8	0	0	0
	SITUATIONAL POSSIBILITY	0	0	0	14	4	0
	ABILITY	0	0	0	1	11	0
	MODAL (OTHER)	0	1	0	2	0	0

Table 4: Inter-annotator agreement for English examples

The number of observed agreements for Spanish examples is 39 (78% of the observations), and the number of agreements expected by chance is 12.7 (25.36% of the observations). The kappa value is 0.705. Therefore, the strength of the agreement is considered to be high.

4 Difficult cases

The agreement rates presented in Section 3.2 are very similar for English and Spanish. This is not surprising given the fact that the modal meanings expressed in both languages are equivalent and they are realized through analogous constructions and lexical items. However, some disagreements and difficult cases were encountered during the annotation of modal tags to our sample corpus of sentences. These can be grouped into two main types:

- a) those cases where the annotator had to decide whether a specific example was modal or not (since the mere presence of a potential modal element does not guarantee the inclusion of modal meanings, as we will see below);
- b) those cases where the annotator had to decide which modality type was conveyed by modal elements (since these are frequently employed to express more than one modal meaning).

With respect to the former, the analysis of problematic as well as unproblematic examples allows us to propose that a prototypical case of modality – one with a high probability of provoking inter-annotator agreement – has the following characteristics:

1. There is a process in the sentence to which modality is applied through the trigger (in bold):

(10) On the one hand, we feel ourselves to be under a greater **obligation** to help those whose misfortunes we have caused. (Multinot: EO_ESSAY_010)

In the example above, the process that is modalized is ‘*help those whose misfortunes we have caused*’. When no process can be identified, human annotators are much more likely to discard the example as a case of modality, as in (11):

(11) Sadly, the United States is not living up to its **obligations**. (Multinot: EO_ESSAY_003)

Of course, between both extremes we find intermediate cases: examples in which a process can be recovered from the preceding text or inferred from context, or nouns which denote processes (*arrival, destruction*) or suggest a process (*the need for an assessment*). It is these cases that have the potential to provoke disagreement between annotators.

2. For all modalities except the *epistemic*, there is a participant in the sentence to which modality is attributed. This participant also plays a semantic role in the process that receives modalization, and it can be explicitly present or recoverable from the context:

(12) Is there any **necessity** to reply to her letter? (Cambridge Dictionary Online)

3. Modality must be the meaning stated in the sentence. If the meaning of modality is presupposed, it is more likely to cause disagreement. Consider the following example: *He has the **potential** to do great things*. This sentence is roughly equivalent to *He **can** do great things*. In both cases, dynamic modality is central to the message the speaker intends to convey. By contrast, in the example below, dynamic modality is presupposed, and the sentence is about what is required to realize that potential.

(12) Realizing the economic **potential** of women requires changes in policies, business practices, and attitudes. (Multinot: EO_ESSAY_004)

Thus, sentences in which the trigger appears as a participant in possession or existential processes (*to have the obligation of, there be no need to*), as the attribute in attributive processes (*it is necessary/obligatory to, he is inclined/prone to*) or as a circumstance (*under the obligation, out of necessity*) cause less disagreement than other constructions.

With respect to those cases where the annotator had to decide among different types of modal meanings, it was found that there were three main types of overlap:

- a) Necessity and Deontic meanings
- b) Possibility and Epistemic meanings
- c) Ability and Situational Possibility

Necessity and the Deontic meanings of Obligation, Prohibition, Absence of Obligation and Recommendation constitute a natural area of overlap, since they are realized by the same lexical items in English and Spanish. The analysis of the cases of agreement and disagreement between human annotators showed that the examples classified unanimously as Necessity exhibited these features in a clear way:

1. The sense of obligation present in the sentence is imposed by nature, not by human norms. It also lacks a sense of morality (what is convenient or correct to do).
2. The sentence has a purely informative function. That is, it does not impose any obligation on the addressee as to a course of action s/he must take.

However, the interference of the human component can be very subtle, as illustrated by example (13) below:

(13) Hatching eggs should be collected soon after lay and maintained at 15-18° C. The eggs **must** not warm to above 20° C. Unless they are being prepared for immediate incubation. (Multinot: EO_EXPE_011)

Here the obligation could be said to have a natural origin. Nevertheless, it is the global purpose of the text – to instruct the reader – that makes annotators hesitate between Necessity and Obligation. A purely scientific text would certainly lead annotators to classify it as Necessity.

Possibility and Epistemic modality are also two closely related concepts. After all, stating that something is possible is often synonymous with conjecturing that something is the case. The problem is more common in Spanish, where the same modal verb (*poder*) covers the meanings of English *can*, *may*, *might* and *could*.

The key factor that seems to favour an Epistemic interpretation over Possibility is that the former concerns the truth of the proposition, and hence creates the implication that the writer ultimately ignores the truth of it, as in the following example.

- (1) Indeed, since the insect head is the seat of some inhibitory nerve centres, it is **possible** that the female improves the male's sexual performance by eating his head. (Multinot: EO_EXPE_003)

The impossibility to determine if the writer is speculating about something s/he ignores or if s/he is describing what is simply possible is a potential cause of disagreement between annotators.

Finally, Ability and Situational Possibility can be problematic categories to distinguish when it is not obvious if the sense of Possibility derives from external conditions or from skills and properties inherent to one participant. In the experiment, the clearest cases of Ability depicted actions that one participant can activate at will and are the result of acquired skills or physical abilities:

- (2) If someone struck a match on the Moon, they could spot the flare. From the tiniest throbs and wobbles of distant stars they **can** infer the size and character and even potential habitability of planets (Multinot: EO_EXPE_001)

By contrast, the most problematic examples involved those cases in which Possibility was attributed to an animate participant, but it was doubtful whether it derived from general enabling conditions or inherent abilities:

- (3) If a strange man touched me I would hit him, and I **can** hit people very hard (Multinot: EO_FICTION_009)

Here it is hard to decide whether 'hitting people' is an ability of the participant 'I', or just an event whose occurrence is compatible with the world as we know it (just like *Summers can be very hot here*). In fact, this seems to be a case of vagueness in the language, rather than a problem of lacking sufficient context to decide which meaning was intended by the writer. In these ambiguous cases, we propose to use the coarser tag in our annotation scheme (Possibility).

5 Conclusion and future work

In this paper we have presented an annotation model for dynamic modality in English and Spanish, within the context of the larger MULTINOT project. The project aims at the creation and annotation of a corpus of English and Spanish texts, original and translated in both directions, for further exploitation in different linguistic and computational contexts. First, we have defined the concept and types of dynamic modality and described a number of problems posed by the study of this category. Secondly, we have set forth an annotation scheme consisting of two interrelated tagsets, which capture the basic functional-semantic meanings and the syntagmatic options that realize these meanings in English and Spanish. This description is followed by a reliability study to validate the proposed scheme. The study was performed on a sample corpus of sentences in English and Spanish, extracted from the MULTINOT corpus and containing lexicogrammatical candidates for the expression of dynamic modality. The results show a high degree of inter-annotator agreement, although some difficult cases and disagreements were encountered during the annotation process. These were basically caused by the following factors: absence of a process expressed by a verb, lack of explicitness of the participant to which modality is attributed, presupposition (instead of assertion) of the modal meaning, and overlap between various modal categories.

We are currently using these findings to develop detailed annotation guidelines for the large-scale annotation of the bilingual texts of the MULTINOT corpus in the upcoming months. Our long-term goal is to produce a high-quality dataset of approximately one hundred bilingual texts enriched with modality features which can be used both in the areas of contrastive linguistics and translation studies, and as training data in a number of computational contexts where modal meanings are essential for the

complete understanding of a text.

References

- Baker, K., M. Bloodgood, B. J. Dorr, C. Callison-Burch, N. W. Filardo, C. Piatko, L. Levin, and S. Miller (2012). Use of modality and negation in semantically informed syntactic MT. *Computational Linguistics* 38(2):1–28.
- Carretero, M. (1995). La pragmática de las expresiones de modalidad epistémica en el inglés hablado. PhD Thesis downloadable at <http://eprints.ucm.es/3368/>
- Collins, P. (2009). *Modals and quasi-modals in English*. Amsterdam: Rodopi.
- Groefsema, M. (1992). 'Can you pass the salt?': A short-circuited implicature. *Lingua* 87: 103–135.
- Halliday, M.A.K., and C.M.I.M. Matthiessen (2014). *Halliday's Introduction to Functional Grammar*. (4th. edition). London and New York: Routledge.
- Hendrickx, I., A. Mendes, and S. Mencarelli (2012). Modality in text: A proposal for corpus annotation. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation - LREC 2012*, Istanbul, May 21-27, 2012. 1805–1812.
- Hermerén, L. (1978). *On modality in English: the study of the semantics of the modals*. Lund: Gleerup.
- Lavid, J., J. Arús, B. DeClerck, and V. Hoste (2015). Creation of a high-quality, register-diversified parallel corpus for linguistic and computational investigations. In *Current work in corpus linguistics: Working with traditionally-conceived corpora and beyond. Selected Papers from the 7th International Conference on Corpus Linguistics (CILC2015)*. Procedia - Social and Behavioral Sciences, Volume 198, 24 July 2015, 249–256.
- Lavid, J., Carretero, M., and Zamorano-Mansilla, J.R. (2016). Contrastive Annotation of Epistemicity in the Multinot Project: Preliminary Steps. Paper accepted for presentation at the *ISA-12, Twelfth Joint ACL - ISO Workshop on Interoperable Semantic Annotation*, held in conjunction with *Language Resources and Evaluation Conference 2016*. Portorov (Slovenia), 28 May 2016.
- Loureiro-Porto, L. (2013) Modal necessity and impersonality in English and Galician. In Marín-Arrese, J. I., M. Carretero, J. Arús Hita and J. van der Auwera (eds.). 2013. *English modality: Core, periphery and evidentiality* Berlin: De Gruyter.
- McShane, M., S. Sergei Nirenburg, and R. Zacharsky (2004). Mood and modality: Out of the theory and into the fray. *Natural Language Engineering*, 19(1):57–89.
- Nuyts, J. (2001) *Epistemic modality, language and conceptualization: A cognitive-pragmatic perspective*. Amsterdam, Philadelphia: John Benjamins.
- Nuyts, J. (2005) The modal confusion: On terminology and the concepts behind it. In A. Klinge and H.H. Müller (eds) *Modality: Studies in form and function*. London: Equinox. 5–38.
- Palmer, Frank R. (1990). [1979]. *Modality and the English modals*. London and New York: Longman.
- Perkins, M.R. (1983) *Modal expressions in English*. London: Frances Pinter.
- Portner, P. (2009) *Modality*. Oxford: Oxford University Press.
- Rescher, N. (1968) *Topics in philosophical logic*. Dordrecht: Reidel.
- Saurí, R., and J. Pustejovsky (2009). Factbank: a corpus annotated with event factuality. *Language Resources and Evaluation* 43(3), 227–268.
- Silva-Corvalán, C. (1995) Contextual conditions for the interpretation of 'poder' and 'deber' in Spanish. In J. Bybee and S. Fleischman (eds) *Modality in grammar and discourse* Amsterdam, Philadelphia: John Benjamins. 67–105.
- Szarvas G, V. Vincze, R. Farkas, J. Csirik (2008). The BioScope corpus: annotation for negation,

uncertainty and their scope in biomedical texts. In BioNLP ACL-2008 workshop

Wärnsby, A. (2006) *(De)coding modality. The case of Must, May, Måste and Kan*. Lund Studies in English 113. Lund: Lund University.

Wiebe, J., T. Wilson, and C. Cardie (2005). Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*, 37(2-3):165–210.

Zamorano-Mansilla, J.R., and M. Carretero (2012). An annotation scheme for dynamic modality in English and Spanish. *Linguistics and the Human Sciences* 6, 282–304.