Guidelines for the annotation of negation in Spanish

Eleonora Guzzi, M. Antònia Martí, Montserrat Nofre, Mariona Taulé

eleonora.guzzi@udc.es, amarti@ub.edu, montsenofre@ub.edu, mtaule@ub.edu
CLiC-Universitat de Barcelona
Universitat de Barcelona Institute for Complex Systems (UBICS)

1. INTRODUCTION

This document presents the guidelines for annotating negation; specifically, the annotation of **negation markers** or **negation cues**, **scope** and **focus**. We describe the annotation scheme adopted and the labels used (tagset). These guidelines are restricted to negation at the syntactic level, i.e. negation in sentences, clauses, constituents and words. We do not take into account lexical or morphological negation.

We identify two different kinds of negative structures. On the one hand, we label structures that contain negation cues even though they do not express negation. Inside this category, we distinguish between traditional comparative sentences and sentences that contain negation markers but actually express an affirmation or a positive statement. On the other hand, we establish a category of true negative structures and distinguish between pure negations and structures that negate an element and, at the same time, establish a contrast with another opposed element.

For the annotation of the focus we take into account the works of Blanco and Moldovan (2014), Altuna et al. (2017), Guzzi et al. (2017) and Francis and Taboada (2017). A complete typology of negation patterns in Spanish is presented in Martí et al. (2016) and Jiménez-Zafra et al. (2017).

Annotation scheme

In this section, we describe the annotation scheme adopted for the annotation of negation in Spanish. The general annotation scheme followed is shown in Figure 1 and the labels used for the annotation are described below.

```xml
<sentence complexity="simple|multiple">
        <neg_structure polarity_modifier="Ø|increment" "Ø"[1]
                        value="neg|contrast|comp|noneg">
                <scope>
                <discount_scope1 disc_scope1_extension="internal|external">
                <discount_scope2 disc_scope2_extension="internal|external">
                <anaphoric_scope anaphora_extension="internal|external">
                <elliptic_scope ellipsis_extension="internal|external">
                        <negexp negtype="Ø|lexical|morphological" "Ø"[1]
                                discid="1n|1c,2n|2c..."
                                negexp_coord="1|2">
                        </negexp>
                        <focus pleonastic_focus="no|yes" "no"[1]>
                        </focus>
                </scope>
        </neg_structure>
</sentence>
```

Figure 1: General annotation scheme (XML structure)

[1] This value is the default value of the attribute.

2.1. <SENTENCE>

The label <sentence> corresponds to a string of words between strong punctuation marks ('.', '?', ';', '!', '…') in which a negative structure can occur.

We only annotate those structures that contain at least one negation marker or negation cue.

When a <sentence> contains a negative structure, it has the attribute <complexity> assigned to it, which has two possible values, ''simple'' or ''multiple'':

– **<complexity=''simple/multiple''>**

We assign the value ''simple'' to the <complexity> attribute when the <sentence> contains only one negative structure (<neg_structure>) (1) and we assign the value ''multiple'' to the <complexity> attribute when the <sentence> contains more than one negative structure (2)[1].

(1) <sentence **complexity=''simple''**> El anterior coche se paró a la media hora de comprarlo <neg_structure> porque <u>no</u> le habían quitado el precinto de seguridad </neg_structure> </sentence> .
'Our previous car stopped half an hour after we bought it [because they had <u>not</u> removed the security seal].'

(2) <sentence **complexity=''multiple''**> <neg_structure>$_1$ para que <u>no</u> les entre polvo </neg_structure>$_1$ o <neg_structure>$_2$ para que <u>no</u> se oxiden </neg_structure>$_2$ </sentence>
'[so that dust does <u>not</u> get in} or [so that they do <u>not</u> rust].'

Complex structures (<sentence complexity=''multiple''>) can be embedded or non-embedded.

Embedded structures (3) and (4) are those in which one negative structure is part of another negative structure in the same <sentence> node.

(3) <sentence **complexity=''multiple''**> <neg_structure>$_1$ <u>no</u> quería pasarme un día entero en el aeropuerto <neg_structure>$_2$ <u>sin</u> poder descansar </neg_structure>$_2$ </neg_structure>$_1$ </sentence> .
'[I did <u>not</u> want to spend the whole day at the airport [<u>without</u> resting]].'

(4) <sentence **complexity=''multiple''**> <neg_structure>$_1$ <u>no</u> tenía culpa <neg_structure>$_2$ de <u>no</u> tenerlo </neg_structure>**$_2$** </neg_structure>$_1$ </sentence>
'[I was <u>not</u> to blame for [<u>not</u> having it]].'

Non-embedded structures are those in which two or more negative structures appear independently in the same <sentence> node (5). They usually correspond to coordinated negative structures (see section 2.4).

(5) <sentence **complexity=''multiple''**> <neg_structure>**$_1$** para que <u>no</u> les entre polvo </neg_structure>**$_1$** o <neg_structure>$_2$ para que <u>no</u> se oxiden </neg_structure>$_2$ </sentence>
'[so that dust does <u>not</u> get in] or [so that they do <u>not</u> rust].'

---

[1] In the examples given hereinafter, we use underlining to mark the negation marker, square brackets to mark the scope and bold type to mark the focus.

2.2. <NEG_STRUCTURE>

The label **<neg_structure>** is assigned to a syntactic structure –corresponding either to a sentence, a clause or a phrase -which contains a negation marker or a negation cue. This label can have two attributes associated with it:

– **<polarity_modifier>**: indicates whether the negative structure contains an element that modifies or nuances its polarity (e.g.: 'no me apetece' / 'no me apetece en absoluto'). This attribute has one possible value: **"increment"** to indicate an increment in the polarity value, as we can see in (6):

> (6) <neg_structure **polarity_modifier="increment"**> No me arrepiento para nada </neg_structure>
> 'I do not regret (it) at all.'

– **<value>:** indicates the meaning expressed by the negative structure. We can distinguish between true negative structures and negative structures that do not express any negative value. In the first category, we can find two possible values, **"neg"** (negative) and **"contrast"**. We can see examples of true negative structures in (7-9).

- **"neg"** indicating negation (7-9):

> (7) <neg_structure **value="neg"**> No recomiendo **el libro** </neg_structure>
> '[I do not recommend **the book**].'

> (8) <neg_structure **value="neg"**> Nadie oía **el dichoso ruído** </neg_structure>
> '[Nobody heard **the annoying noise**].'

> (9) Ha llegado a lavar un edredón nórdico <neg_structure **value="neg"**> sin **problema** </neg_structure>.
> 'I've even washed a nordic quilt [without **trouble**].'

The value **"neg"** can be associated with the <polarity_modifier> attribute, as we can see in (10), where the negation marker 'nada' expresses an <increment>:

> (10) <neg_structure **value="neg" polarity_modifier="increment"**> El final del libro no te aporta **nada** </neg_structure>.
> '[The end of the book does not add **anything**].'

- **"contrast"** indicating contrast or opposition between terms (11-13):

> (11) <neg_structure **value="contrast"**> No vinieron **2 soldados,**</neg_structure> sino 6.
> '6 soldiers came, [not **2 soldiers**].'

> (12) <neg_structure **value="contrast"**> El islam no es **una religión,**</neg_structure> sino un sistema político.
> '[Islam is not **a religion**], but a political system.'

(13) <neg_structure **value="contrast"**> <u>No</u> pago mis impuestos **para los bancos** </neg_structure>, sino para pagar las pensiones de mis padres.
   '[I do<u>n't</u> pay my taxes **for the banks]**, but to pay for my parents' public pensions.'

In these cases, the negative structure corresponds to the first part of the contrastive sentence. The scope will include only the negated part of the sentence, with its corresponding focus and negation marker or cue (see Section 2.3.4).

In the second category, negative structures that do not express negation, we distinguish between comparative structures and structures that should not be considered negative if we take into account their semantic and pragmatic value. For the first class, we use the value **"comp"** and for the latter, the value **"noneg"**:

- **"comp"** expressing either a comparison or a condition of inequality between terms (14-15) or a superlative degree (16-17). In all these cases, only terms such as 'no', 'nada', 'ni', 'ningún' are considered negative expressions. Therefore, comparative elements such as 'tan', 'tanto', 'como' are not labelled as <negexp>.

  Note how these structures, despite having a negation marker, express an affirmative meaning. Thus, in (14), the writer is not saying that he/she does not like the exterior, but that he/she likes it less than in other brands. In (15) the engine is less powerful than it should be, but the reader could infer that the engine is actually powerful. In (16) the writes is stating that the public system is the best possible system and in (17) that the solution they propose is the best one.

  Lastly, these negative structures are not labelled with the tag <scope> or <focus>, because they do not in fact express negation:

(14) <neg_structure **value= "comp "** > su exterior <u>no</u> me gusta tanto como el de otras marcas </neg_structure>
   'I do<u>n't</u> like the outside as much as other brands.'
(15) <neg_structure **value="comp"**> El motor no es todo lo potente que debería </neg_structure>
   'The engine is <u>not</u> quite as powerful as it should be.'
(16) <neg_structure **value="comp"**> <u>No</u> hay <u>nada</u> mejor que el sistema de pensiones públicas </neg_structure>
   'There is <u>nothing</u> better than the public pensions system.'
(17) <neg_structure **value="comp"**> <u>No</u> hay <u>ninguna</u> solución mejor que la propuesta. </neg_structure>
   'There is <u>no</u> better solution than the one we propose'.

- **"noneg"** indicating structures that contain a negation marker but which do not negate (18-21):

(18) El coche lo compré para viajar, <neg_structure **value="noneg"**> <u>no</u>? </neg_structure>
'I bought this car for travelling, did<u>n't</u> I?'

(19)<neg_structure **value="neg"**> No pienso irme <neg_structure **value="noneg"**> hasta que <u>no</u> vengas. </neg_structure></neg_structure>
'I'm not leaving until he/she comes (literally '<u>doesn't</u> come').'

(20)<neg_structure **value="noneg"**> Esta mujer es <u>ni más ni menos</u> que madame Bertholdi. </neg_structure>
'This woman is <u>no more and no less</u> other than Madame Bertholdi.'

(21)<neg_structure><neg_structure **value="noneg"**> <u>Hasta que no</u> coloca la ropa</neg_structure> no centrifuga. </neg_structure>
'Until you put the clothes in the spin-dryer it does<u>n't</u> start (literally 'until you do<u>n't</u> put in').

Regarding negative structures with the markers 'no_solo'...'sino_también', we use the value **"noneg"**. We consider them "noneg" because they are not negating anything or contrasting two elements but adding them. These sentences are not labelled with the tags <scope> or <focus> (22-23):

(22)<neg_structure value=**"noneg"**> <u>No solo</u> quieren dinero, </neg_structure> sino_también tener más poder.
'They do <u>not_only</u> want money, but_also more power.'

(23)<neg_structure value= **"noneg"**> es un coche pensado <u>no solo</u> para su uso</neg_structure>, sino para su disfrute
'This car is designed <u>not_only</u> to be used, but_also to be enjoyed'.

2.3. <SCOPE>

The attribute <**scope**> delimits the part of the negative structure that is within the scope of the negation. The scope includes all the words affected by negation. When annotating the scope, the maximum range of words affected by the negation must be selected (Vincze et al. 2008; Konstantinova et al. 2012; Francis and Taboada, 2017). Unlike these authors, we include the negation marker or cue <negexp> within the scope. Following Guzzi et al. (2017), Francis and Taboada (2017) and Altuna et al. (2017) we annotate the focus and include it within the scope.

In our approach to the annotation of negation, the subject is always included within the <scope> when the word directly affected by negation is the verb of the sentence

(24)<**scope**>España <u>no</u> reacciona **ante estas barbaridades del gobierno**</scope>.
'[Spain does<u>n't</u> react **to these government barbarities**.]'

- <discontinuous scope> and <elliptical scope>:

There are sentences in which the scope is discontinuous, that is, part of the scope is outside of the negative structure containing the negation marker and the focus (25-26).

(25)<discontinuous_scope1 **discont_scope1_extension="internal"**> **Yo cotizo encantado** </discontinuous_scope1> para los jubilados de hoy, <discontinuous_scope2 **discont_scope2_extension="internal"**> <u>no</u> **para beneficio de los putos bancos** </discontuous_scope2>.
'[I'm glad to contribute]₁ for today's pensioners,[<u>not</u> **to benefit banks**.]₂'

(26)<discountinuous_scope1 **discont_scope1_extension="internal"**>**Las pensiones se asignan**</discountinuous_scope1>por individuo y <discountinuous_scope2 **discont_scope2_extensions="internal"**><u>no</u> **por territorio**</discount_scope2>
'[**Pensions are assigned on**]₁ an individual basis, [<u>not</u> **a territorial basis**.]₂'

In other sentences, the scope and the focus are located outside the negative structure containing the negation marker. We consider that the focus is elliptical in these cases.

(27) Les da igual si <anaphoric_scope **anaphora_extension="internal"**>llegan **vivos** </anaphoric_scope> o <elliptic_scope **ellipsis_extension="internal"**><u>no</u> </elliptic_scope>.

'They don't care if [they make it **alive**]₍anaphoric_scope₎ or [<u>not</u>.]₍elliptic_scope₎'

Both discontinuous and elliptical scopes have an attribute for indicating the extent of the scope. In the case of discontinuous scope we use the attribute <disc_scope_extension="internal"|"external">.

In case of discontinuous scope, we use the value "internal" when the two parts of the scope are in the same sentence (25-26). We use the value "external" when the two parts of the scope are included in two different sentences (and are therefore independent from the syntactic point of view) (28).

(28) <discontinuous_scope **disc_scope1_extension="external"**>**Me refiero** </discontinuous_scope1> a los niños.
<discontinuous_scope2 **disc_scope2_extension="external"**><u>No</u> **a los adultos**</discontinuous_scope2>, el que ya está adoctrinado es insalvable.
'[I mean]₁ the children.
[Not adults,]₂ the ones who are already indoctrinated are unsalvable.'

In case of elliptical scope, we use the value "internal" when the two parts of the scope are in the same sentence (27). We use the value "external" when the two parts of the scope are included in two different sentences (and are therefore independent from the syntactic point of view) (29).

(29) <anaphoric_scope **anaphora_extension="external"**>**Planes de pensiones privados**</anaphoric_scope>? <elliptic_scope **ellipsis_extension="external"**><u>No</u>, gracias...</elliptic_scope> No me apetece perder dinero

'[**Private pensions plans**?]₍anaphoric_scope₎ [<u>No</u>, thanks...]₍elliptic_scope₎ I don't feel like losing my money'.

2.3.1. <negexp>

This label is used for the word or words that express negation (negation markers or negation cues). Negation can be expressed by means of very different mechanisms, with the use of syntactically independent negation particles being the most common (e.g., *no, nunca, nadie*)(30), but it can also be expressed by means of prefixes (e.g., *impensable, ilegible*) (32) and idioms (e.g., *en la vida, en absoluto*) (31). So far we have only annotated syntactic negation, but we have created an attribute to mark the other ways to express negation (<negtype="∅|lexical"|"morphological">). Therefore, by default, the attribute <negtype> is omitted in the case of syntactic negation.

(30) **<negexp>**No</negexp> hubo **acuerdo**.
   '[There was no **agreement**.]'
(31) Me importa **<negexp negtype="lexical">**un pimiento</negexp> **lo que piensen de mí**.
   '[I don't give a damn what they think of me].'
(32) Para   que   los   cristianos   no   me   vean   **<negexp
   negtype="morphological">**inmoral</negexp>.
   '[So that Christians won't see me as immoral].

We annotate negative expressions that belong to different gramatical categories: adverbs (*no, jamás, nunca, tampoco, nada*) or adverbial phrases *(en absoluto, en ningún momento)*; pronouns (*nada, nadie, ninguno, nunca*); conjunctions (*ni*); prepositions (*sin*), indefinite determiners (*ningún, ninguna*) and some idioms that include a negation marker (*ni una palabra, en ningún momento, ni siquiera*).

2.3.2. <discid>

The negative expression can include one or more than one negative element. In the latter case, the elements can be continuous or discontinuous and the second negative element usually nuances the first one, as in (33).

(33)<neg_structure   polarity_modifier="**increment**"> No **me arrepiento** para nada
   </neg_structure>
   '[I do not **regret** (it) at all.]'

Within <negexp> we use the attribute <discid> for tagging the discontinuous elements in the negative expression.

When negation markers are **discontinuous**, we identify the negative elements by means of two <negexp> labels, each of them with the attribute <discid> (discontinuity id). The value of <discid> is represented both numerically (1 in the example below), which indicates the numerical order of the discontinuous negative elements in a sentence, and by a letter 'n' and 'c', where 'n' and 'c' indicate the first and second elements of the negation respectively, as we can see in (34-35):

(34)**El coche** <negexp **discid="1n"**> no </negexp> frena <negexp **discid="1c"**> en absoluto
   </negexp>
   '[**The car** does not$_1$ brake at all$_2$.]'

(35) **Eso** <negexp **discid="1n"**>no</negexp> sirve <negexp **discid="1c"**>de nada.</negexp>
   '[**That** doesn't help.]'

2.3.3. <focus>

The basic linguistic assumption on which we base our proposal is the Huddleston and Pullum (2002) approach, in which the focus of negation is the part of the scope that is most prominently or explicitly negated.

Our annotation proposal assumes three criteria for identifying focus:

First, we consider the discourse context criterion: we take into account as discourse context the whole comment that contains the negations structure. We do not consider the previous and the following comments because they are not necessarily connected to the comment under analysis: the temporal thread does not guarantee that there exists a connection between a comment and the ones preceding and following it. Therefore, we take into account the inter-sentential relationships within the comments. All the comments refer to one online news article that can be considered their referential world. This news article also contains an important part of the pragmatic world knowledge necessary to understand the content of the comments. Therefore, we take into account this information in the annotation process.

Second, we consider the obliquity criterion. We assume that the most oblique argument in a sentence or in a clause is the most plausible candidate to be the focus, with the adjuncts the most oblique of the arguments. The underlying idea is that negation affects the most specific (oblique) information, otherwise this information would not be explicitly stated, and this information is expressed because it is what we want to negate. We annotate the whole argument not part of it.

Third, we consider the criterion of implicit positive meaning (Blanco and Moldovan 2011, 2014), when possible: "(the focus of negation is) the element of the scope that is intended to be interpreted as false to make the overall negative true, therefore a negated statement can carry a positive implicit meaning".

We distinguish between the explicit and implicit focus. The explicit focus is expressed by means of formal markers such as displacement and explicit pronominal subjects. We define the implicit focus to be when there are no formal markers for its identification. In this case, we apply the most oblique argument criterion (Guzzi et al. (2017) and Francis and Taboada (2017)), as long as the context does not give other information. Taking into account the oblique criterion, we distinguish between arguments and adjuncts. When in a negation structure there is an adjunct we consider it to be the most oblique element and, therefore, the focus. If there is more than one adjunct, we consider manner to be the most oblique argument followed by place, and time, although in this specific case we are considering the possibility of accepting more than one focus in a future updated version of the corpus.

Regarding the arguments, the most oblique will be the indirect object, followed by the prepositional object, the direct object, and the least oblique will be the subject. When the negated sentence contains only one verb, it will be the focus.

The negation structure corresponds either to a sentence, a clause or a phrase. In our approach, the focus is always included in the scope and corresponds to:

- a verb form (36),

(36) <scope><u>No</u> **<focus>pasará</focus>**</scope>, ya ha pasado.
    '[It wo<u>n't</u> **happen**], it's already happened.'

- an argument (37)

(37)Dice que <scope><u>no</u> vendrá **<focus>Luisa</focus>**</scope>**.**
    'He/She says that [it is **Luisa** who wo<u>n't</u> come].'

- or an adjunct (38).

(38)<scope><u>No</u> lo consideran **<focus>así</focus>**</scope>.
    '[They do<u>n't</u> consider it to be **that way**].'

Arguments and adjuncts can be syntactically realized as a phrase (37) and (38) or as a clause (39-40). This is in accordance with Blanco and Moldovan's (2011, 2014) proposal, in which the focus is always the full text of a semantic argument (or adjunct).

(39) <scope><u>No</u> tiene problemas **<focus>para volver a entrar en el grupo</focus>**</scope>.
    '[He/She doesn<u>'t</u> object **to rejoining the group**.]
(40)<scope><u>No</u> quiere **<focus>escribir un artículo de política</focus>**</scope>**.**
    '[He/She does<u>n't</u> want **to write an article about politics**.]'

Note that in some examples more than one element are good candidates to be considered as the focus (41a, 41b):

(41) (a) <scope>Eso <u>no</u> debería ser posible **<focus>con una buena vigilancia</focus>** aplicando medidas drásticas</scope>.
    '[That would<u>n't</u> happen **if there was adequate supervision**, applying extreme measures].'
(b) <scope>Eso <u>no</u> debería ser posible con una buena vigilancia **<focus>aplicando medidas drásticas</focus>**</scope>.
    '[That would<u>n't</u> happen if there was adequate supervision, **applying extreme measures**].'

We are considering the possibility of accepting either more than one focus when such ambiguity occurs or considering the last element in the sentence to be the focus ('aplicando medidas drásticas'). At the moment, in these cases, the selection of the focus depends on the criterion of the individual annotator.

2.3.3.1. Explicit focus

We identify five types of explicit focus: displacements, explicit pronominal subjects, contrastive constructions, reinforcement and typographic clues.

- Displacements

Displacement is a focalization mechanism consisting of moving the focused element into a marked position, usually fronted.

> (42) Eso sí, tenemos el nivel de alquiler de Europa, somos 'la polla'. <scope>Y **<displacement_focus>de los sueldos europeos (de los que no disfrutamos los españoles)**</displacement_focus> <u>no</u> dicen <u>nada</u>?.</scope>
> 'That's right, we have the same rent prices as Europe, we're the 'bee's knees'. [And what **about European salaries** do<u>n't</u> they say <u>anything</u> about them?].'

The noun complement in (42) 'de los sueldos europeos (…)' is an example of leftward displacement of a noun complement to emphasize this constituent in order to be considered as the focus.

Another type of displacement is the **emphatic subject**, that is, the subject is displaced to a postverbal position (43).

> (43) <scope> <u>No</u> reparte números, <displacement_focus> **una servidora** </displacement_focus> </scope>.
> '[(She) does<u>n't</u> hand out numbers, **yours truly**.]'

A specific type of displacement is **pleonastic focus**. This happens when an argument is expressed twice in a sentence, once in a displaced position 'A su sobrino' (44a) or 'A mí' (44b), and then again as a pronoun 'le' (44a) and 'me' in (44b) -inside the sentence and before the verb. In order to distinguish between these two focus expressions, we tag the former as pleonastic focus with the attribute <pleonastic_focus="yes"> and the latter (the pronoun) simply as <focus>.

> (44)(a) <scope>**<focus pleonastic_focus="yes">A su sobrino**</focus>, <u>no</u> <focus>**le**</focus> había tocado un piso de protección oficial</scope>.
> '[**His/Her nephew...**pleonastic , **he/she** has<u>n't</u> been given an official protection flat].'
> (b) <scope><focus pleonastic_focus="yes">**A mi**</focus>, <u>no</u> <focus>**me**</focus> parece mal introducir un poco de mentalidad anglosajona.<scope>'
> '[**I...**pleonastic , **I** do<u>n't</u> think it's a bad idea to introduce a little Anglo-Saxon mentality].

We tag as displaced those verb arguments that have been displaced to the beginning of the sentence. We do not consider adverbial complements (adjuncts) that express time, location and manner to be displaced focus when they appear in a preverbal position because their position in a sentence is free in Spanish.

> (45) <scope><focus>**En España**</focus>, <u>nunca</u> van a triunfar los neoliberales</scope>.
> '[**In Spain**, the neoliberals will <u>never</u> triumph].'

- Explicit Pronominal subject

Spanish is a null subject (pro-dop) language and the subject is not usually explicitly expressed. That is why when a pronominal subject is explicit in a sentence it is because the writer wants to highlight the communicative role of the pronoun in the sentence. We consider these explicit pronominal subjects to be the focus of the negative structure when they appear with a verb (46) or in a contrastive construction (47), where the writer contrasts 'they' with 'we':

(46) El agnóstico es el que dice: <scope> **<focus> Yo** </focus> <u>no</u> creo </scope>, pero no vaya a ser...
   'The agnostic is the one who says: [**I** do<u>n't</u> believe], but it could be...'

(47) <scope> Si <focus> **ellos** </focus> <u>no</u> hacen <u>nada</u> </scope>, nosotros tampoco.
   '[If **they** do<u>n't</u> do <u>anything</u>], neither do we.'

- Contrastive constructions

Contrastive constructions, introduced by 'pero', 'no obstante', 'sino', etc., help in the detection of the focus of negation as they express the element which is in contrast ((48) and (49)) with the focus. The contrastive construction in (48) marks the focus 'Pedro' by introducing the alternative object ('Juan'), which is in contrast. Whereas in (49) the focus is the verb 'vino' because the element in contrast is the verb 'llamó'.

(48) <scope><u>No</u> vino <focus> **Pedro** </focus> </scope>, sino Juan.
   '[It was<u>n't</u> **Pedro** who came], but Juan.'

(49) <scope> <u>No</u> <focus> **vino** </focus> Pedro </scope>, pero llamó.
   '[Pedro did<u>n't</u> **come**] but he did call.'

In conclusion, the detection of the focus in the negative structures labelled as "contrast" should be very straightforward, as the element after the contrastive expression ("sino", "más bien" or even a comma) should help to establish the alternative object and, therefore, the focus of the negation.

- Reinforcement of negation

Reinforcement is another explicit mechanism for marking the focus of negation. Reinforcements are negative constructions that contain two or more negation markers or cues. They usually consist of the adverb 'no' and a second, usually discontinuous, negation marker (50-51):

(50) <scope> <negexp **discid="1n"**><u>No</u></negexp> tiene **<focus>** <negexp **discid="1c"**><u>**ningún**</u></negexp>**motivo**</focus></scope>.
   '[He/She does<u>n't</u> have **no reason** (literally).]'

(51) <scope> <negexp **discid="1n"**><u>No</u> </negexp> ha comprado **<focus><negexp discid="1c">**<u>nada</u></negexp></focus></scope>.
   '[He/She has<u>n't</u> bought **anything** (**nothing**, literally).]'

Note that it is not uncommon that the second part of a discontinuous negative expression is, at the same time, the focus of the negative structure, as in (51). In those cases, we use first the label <focus>, in a higher position, and then the label <negexp **discid="1c"**>.

- **Typographic clues**

Typographic clues are considered explicit markers and a strategy to emphasize an element of the negation structure, the focus of negation. They include uppercase letter, bold, underlined elements and italics. The element affected by these typographic changes is the focus of negation.

> (52) <u>No</u> parecerían tan listos <scope>si no tuvieran **<focus> prácticamente TODOS los medios de información** </focus> bajo su control </scope>.
> 'They would<u>n't</u> seem so smart [if they did<u>n't</u> control **ALL the media**.]'

In (52), we assign the focus to 'prácticamente TODOS los medios de comunicación' because of the presence of the typographic clue. Otherwise, the focus would be 'bajo su control' because it is the most oblique argument.

2.3.3.2. Implicit focus

We use the term implicit focus of negation to refer to those cases in which there are no formal markers that allow for its identification. In this case, we assume that the most oblique argument or adjunct in a sentence is the most plausible candidate to be the focus. The underlying idea is that negation affects the most specific (oblique) information, otherwise this information would not be explicitly stated, and this information is expressed because it is what we want to negate. In implicit focus, we distinguish between constituent focus and sentence focus according to whether the scope affects a constituent or the whole sentence.

- **Constituent Focus**

In this case, the focus is an element within a constituent (a phrase), where the constituent is the scope (53) and (54):

> (53) Un problema <scope><u>no</u> **<focus>muy preocupante**</focus></scope>
> '[A **not-too-worrying**] problem.'
> (54) Coca-cola <scope><u>sin</u> **<focus> cafeína** </focus></scope>.
> 'Coca-cola [<u>without</u> **caffeine**.]'

- **Sentential focus**

We distinguish two types of sentence focus: a) when the focus is an argument and b) when the focus is an adjunct. We describe how to represent the focus when an argument or an adjunct is expressed by a subordinate clause in subsection c) below.

a) Argument as focus of negation

In the case of **intransitive verbs** without adjuncts, the focus can be the verb or the subject (external argument) depending on the context (55-56). The meaning is often ambiguous, and, in these cases, we apply the oblique criterion and mark the explicit subject as the focus, when the context does not help in its identification.

> (55) <scope><u>No</u> **<focus>pasará**</focus></scope>, ya ha pasado.

`[It won't **happen**], it's already happened.'

(56)<scope>**<focus>Vuestros dioses</focus>** no existen.</scope>
    '[**Your gods** don't exist].'

In the case of **existential verbs**, the focus is the internal argument, that is, the existential subject, because the verb is lexically empty (Morante, Schrauwen and Daelemans 2011) (57):

(57)<scope>No hay **<focus>agua</focus>**</scope>.
    '[There isn't **water**.]'

In the case of verbs with two arguments, that is, **transitive verbs** (58), **copulative verbs** (59) and **verbs with a prepositional object** (60), the focus is the direct object, the attribute and the prepositional complement respectively.

(58)<scope>No se especifica **<focus>el precio</focus>**</scope>.
    '[**The price** is not especified.]'
(59) <scope>La propuesta no es **<focus>una novedad</focus>**</scope>.
    '[The proposal is not **an innovation**.]'
(60)<scope>No trato **<focus>con los asesinos</focus>**</scope>.
    '[I don't deal **with murderers**.]'

In the case of **light verbs** such as 'haber', 'dar', 'hacer', the focus is the corresponding complement (61).

(61)<scope>No ha hecho **<focus>la limpieza</focus>**</scope>.
    '[He/She hasn't done **the cleaning**.]'

Ditransitive verbs require three arguments: subject, direct object and indirect object/ prepositional complement. The criterion applied in these cases is to consider the most oblique argument as the focus (62):

(62)<scope>María no regaló la camisa **<focus>a Pedro</focus>**</scope>.
    '[María didn't give the shirt **to Pedro**.]'

In the case of **periphrastical verbs**, we apply the same criterion as for verbs with one, two or three arguments, taking into account the argument structure of the verb in the non-finite form (gerund, past participle or infinitive) (63).

(63)<scope>La rabia no va a vencer **<focus>al odio</focus>**</scope>.
    '[Rage won't defeat **hate**.]'

When the focused verbal argument has a complement, the focus is the whole argument, including the head and its complements. The head of an argument can be:

- a noun (64)

> (64)<scope>No ha hecho **<focus>una descripción de la operación**</focus></scope>.
> '[He/She has<u>n't</u> given **a description of the operation**.]'

- an adjective (65)

> (65) Está claro <scope>que <u>no</u> es **<focus>tan fácil**</focus></scope>.
> 'It's clear, [that it is<u>n't</u> **that easy**]'.

- or and adverb (66)

> (66) Esperemos <scope>que a los españoles <u>no</u> nos etiqueten <**focus>tan libremente**</focus>cuando hacemos turismo</scope>.
> 'Let's hope [that Spanish tourists do<u>n't</u> get labelled **so freely** when we travel abroad].

b) Adjuncts as focus of negation

Since adjuncts are optional, their presence in negative structures denote that they carry relevant information (67-69) and constitute the focus of negation.

> (67)<scope>Las pruebas <u>no</u> han proporcionado, **<focus>hasta el momento**</focus>, resultados apreciables</scope>.
> '[**Till now**, the tests have <u>not</u> provided appreciable results.]'
> (68)<scope><u>No</u> quiere comer **<focus>aquí**</focus></scope>.
> '[He/She does<u>n't</u> want to eat **here**.]'
> (69)<scope><u>No</u> puede explicarse **<focus>en pocas palabras**</focus></scope>.
> '[It can<u>not</u> be explained **in few words**.]'

It is worth nothing that the restrictive adverbs such as 'solo', 'solamente', 'únicamente' are the focus of negation. In this case, what is negated is the restriction denoted by the adverbs. In (70) what is negated is that (s)he is devoted exclusively to teaching:

> (70)<scope>No se dedicará **<focus>solo**</focus> a la enseñanza</scope>.
> '[He/She wo<u>n't</u> devote him/herself **only** to education.]'

c) Focus and subordinate clauses

When the most oblique argument is a subordinate clause (a nominal or an adverbial clause) the focus is the whole clause.
Sentences (71-74) are examples in which the focus is a nominal subordinated clause with different syntactic functions: subject (71), attribute (72), direct object (73) and prepositional object (74).

> (71) <scope>Lo malo <u>no</u> es **<focus>que te guste**</focus></scope> sino que dejes que afecte a tu vida.
> `[The bad thing is <u>not</u> **that you like it**] but that you let it affect your life.'

(72) Esto parece que <scope><u>no</u> es **<focus>lo que desean nuestros amados líderes**</focus></scope>].

`[This does<u>n't</u> seem to be **what our beloved leaders desire**].'

(73) <scope><u>No</u> sé **<focus>si ves la diferencia**</focus></scope>.

`[I do<u>n't</u> know **if you can see the difference**]'

(74) <scope>La gente <u>no</u> se queja **<focus>de que hagas horas extras**</focus></scope>.

`[People do<u>n't</u> complain **about you doing overtime**].'

2.4. <NEG_COORD>

We assign the label <neg_coord> to a coordination between two or more negative structures. Inside coordinated negative structures, the negative expressions have the attribute <negexp_coord> attached to the label <negexp>. We assign the attribute "1" to the first negative expression and the attribute "2" to the second and subsequent negative expressions.

We identify two different types of coordinated structures.

a) The first type are those structures in which each negative structure has its own scope, negation marker and focus. Thus, under a wider label <neg_coord>, each negative structure has its own <neg_structure>, <scope>, <negexp> and <focus> (75-76).

(75) **<neg_coord>** <neg_structure>$_1$ <scope><negexp **negexp_coord="1"**><u>No</u></negexp> pidas <focus>**a quien pidió** </focus></scope></neg_structure>$_1$ <neg_structure>$_2$ <scope> <negexp **negexp_coord="2"**> <u>ni</u> </negexp> sirvas <focus> **a quien sirvió** </focus></scope></neg_structure>$_2$ </neg_coord>.

'[Do<u>n't</u> beg someone **who begged**] and [do<u>n't</u> serve someone **who served**].'

(76) **<neg_coord>**<neg_structure>$_1$<scope><negexp **negexp_coord="1"**><u>Ni</u></negexp> <focus>**nihilismo**</focus></scope></neg_structure>$_1$ <neg_structure>$_2$ <negexp **negexp_coord="2"**><u>ni</u></negexp><focus>**todos iguales**</focus> </scope> </neg_structure>$_2$ </neg_coord>.

'[<u>Neither</u> **nihilism**] [<u>nor</u> **everybody is equal**].'

b) The second type are coordinated negative structures in which two or more structures share at least part of the scope. We use the label <scope> within the label <neg_coord>. Inside this wide scope, we mark the different negative structures as usual but without the tag <scope> (77-78):

(77) **<neg_coord> <scope>** <neg_structure>$_1$ Sánchez parece que <negexp negexp_coord="1"><u>no</u></negexp> <focus>**aprende** </focus></neg_structure>$_1$ <neg_structure>$_2$ <negexp negexp_coord="2"><u>ni</u> </negexp><focus> **aprenderá** </focus></neg_structure>$_2$ </scope> </neg_coord>.

'[It seems that Sánchez does<u>n't</u> **learn** and wo<u>n't</u> **learn**].'

(78) **<neg_coord> <scope>** <neg_structure>$_1$ La guerra <negexp negexp_coord="1"><u>no</u> </negexp>terminó <focus> **con la victoria de los fascistas** </focus></neg_structure>$_1$ <neg_structure>$_2$<negexp negexp_coord="2"> <u>ni</u> </negexp><focus>**con los años de gobierno del tío paco** </focus></neg_structure>$_2$ <neg_structure>$_3$ <negexp

negexp_coord="2">ni </negexp><focus>**con la tan modélica transición** </focus></neg_structure>3 </scope> </neg_coord>.

'[The war did<u>n't</u> finish **with the victory of the fascists**, <u>nor</u> **with uncle Paco's government**, <u>nor</u> **with the supposedly exemplary transition]**.'

In some cases, the focus is shared by two or more coordinated negative structures (79). Note that we label only one focus for both negative structures and only within one of the coordinated negative structures.

(79) Solo podemos educar a la sociedad para que denuncie y **<neg_coord> <scope>** <neg_structure>1 <negexp negexp_coord="1"><u>no</u> </negexp>tolere </neg_structure>1 <neg_structure>2 <negexp negexp_coord="2"><u>ni</u> </negexp>apoye **<focus> esos movimientos** <\focus> </neg_structure>2 </scope> </neg_coord>.
'We can only educate society to make them denounce and [<u>neither</u> tolerate <u>nor</u> support **these movements**].'

In some cases, the corresponding negative expression is not expressed in the sentence. We should therefore understand that these are examples of an elliptic expression of negation. In (80) a negative expression such as 'ni' must be understood as also coming before 'vivir dignamente sin ayudas sociales':

(80) No es muy inteligente que sigan llegando personas <neg_coord> <scope> <neg_structure>1 que<negexp negexp_coord="1"> <u>no</u> </negexp>podrán <focus> **acceder a un trabajo** </focus></neg_structure>1, <neg_structure>2 <focus>**vivir dignamente sin ayudas sociales** </focus></neg_structure>2 <neg_structure>3 <negexp negexp_coord="2">ni </negexp>**integrarse activamente en la sociedad** </neg_structure>3 </scope> </neg_coord>.
'It's not very smart that more people [who wo<u>n't</u> be able **to find a job**, <u>nor</u> have **a good life** <u>nor</u> **integrate in society keep arriving**].'

In example (81), we find a coordinated negative structure embedded in another coordinated structure.

(81) **<neg_coord>1** <scope> <neg_structure>1 Los gobiernos sumisos a la UE<negexp negexp_coord="1"> <u>no</u> </negexp>se atreven </neg_structure>1 **<neg_coord>2** <scope><neg_structure>2 <negexp negexp_coord="2"><u>ni</u> </negexp>quieren saber </neg_structure>2 <neg_structure>3 <negexp negexp_coord="2">**<u>ni</u>** </negexp>actuar <focus>**en este tema**</focus></scope> </neg_structure>3 </neg_coord>2 </scope> </neg_coord>1.
'[The governments which are submissive to the EU do<u>n't</u> dare [<u>nor</u> want to know <u>nor</u> want do something **about this issue**]].'

References

Altuna, B., Minard, A. and Speranza M. (2017). *Guidelines: Negation in Italian*.

Altuna, B., Minard, A. and Speranza M. (2017). The Scope and Focus of Negation: A Complete Annotation Framework for Italian. *Proceedings of EACL 2017, SEMBEAR workshop*, to appear.

Blanco, E. & Moldovan, D. (2011) Semantic Representation of Negation Using Focus Detection. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics pp. 581–589. Association for Computational Linguistics.

Blanco, E. & Moldovan, D. (2014) Retrieving implicit positive meaning from negated statements. Natural Language Engineering 20 (4): 501–535. Cambridge: Cambridge Univer- sity Press.

Francis, E. and Taboada, M. (2017). Guidelines for Annotating negation, Scope and Focus. Discourse processing La., Simon Fraser University, Canada.

Guzzi, E., Taulé, M. and Martí M. A. (2017). Criterios para la detección del foco de la negación en español. *Workshop NEGES-2017*, SEPLN-2017, Murcia, Spain.

Huddleston and Pullum (2002)

Jiménez-Zafra, S., M. Taulé, M. T. Martín-Valdivia, M. A. Martí, L. A. Ureña López (2017). SFU ReviewSP-NEG: a Spanish corpus annotated with negation for sentiment analysis a typology of negation patterns, *Language, Resources and Evaluation*, Springer Science+Business Media Dordrecht. DOI: 10.1007/s10579-017-9391-x.

Konstantinova, N., S. C de Sousa, N. P. Díaz, N. P. Cruz, M. J. Maña, M. Taboada and R. Mitkov. 2012. A review corpus annotated for negation, speculation and their scope. *Proceedings of the 8th International Conference on Language Resources and Evaluation* (LREC'12), páginas 3190-3195, Turkey.

Martí, M.A., M. T. Martín-Valdivia, M. Taulé, S. Jiménez-Zafra, M. Nofre y L. Marsó. (2016). La negación en español: análisis y tipología de patrones de negación. *Procesamiento del Lenguaje Natural*, Volume 57: 41-48.

Morante, R., S. Schrauwen y W. Daelemans. (2011). Corpus-based approaches to processing the scope of negation cues: an evaluation of the state of the art. *Proceedings of the Ninth International Conference on Computational Semantics* (IWCS 2011). Johan Bos and Stephen Pulman (editors), pp. 350-354, Oxford, UK.

Vincze, V., Szarvas G., Farkas R., Móra G. and Csirik J. 2008. The BioScope corpus: biomedical texts annotated for uncertainty, negation and their scopes. *BMC Bioinformatics*, 9:1-9.