THE SCOPE AND FOCUS OF NEGATION:
A Complete Annotation Framework for Italian

*Begoña Altuna, **Anne-Lyse Minard and **Manuela Speranza
April 4, 2017

*University of the Basque Country, **Fondazione Bruno Kessler
Introduction

Negation in Italian

Negation cue, Scope and Focus

Annotating negation in texts

Discussion
INTRODUCTION
Negation: Negation is a linguistic phenomenon that inverts the truth value of the proposition it is applied to (Martí et al., 2016).
Example

Syria’s women’s team has **not** played a competitive international since 2011.
Elements of negation

**Negation cue:** lexical and syntactic elements that express negation.

**Example**

Syria’s women’s team has *not* played a competitive international since 2011.
Example

Syria’s women’s team has not played a competitive international since 2011.
**Scope**: the extent of the text affected by the cue (Morante *et al.*, 2011) or the section of text expressing the proposition the truth value of which is inverted by the negation.

**Example**

Syria’s women’s team has *not* played a competitive international since 2011.
Example

Syria’s women’s team has \textcolor{red}{not} played a competitive international since 2011.
Elements of negation

Focus: part of the scope that is most prominently or explicitly negated (Huddleston and Pullum, 2002).

Example

Syria’s women’s team has **not** played a competitive international since 2011.
Relevance of negation annotation

Negation information is relevant in many NLP tasks:

- Event factuality resolution
- Discrimination of the entities taking part in an event
- Reversion of the polarity in sentiment analysis
- Discourse relation extraction
State of the art

Negation mark-up schemes

- BioScope guidelines (Vincze et al., 2008)
- ConanDoyle-neg guidelines (Morante et al., 2011)

Corpora

- BioScope (Vincze et al., 2008)
- ConanDoyle-neg (Morante and Daelemans, 2012)
- SFU Review Corpus (Konstantinova et al., 2012) and its Spanish version (Martí et al., 2016)
- Product Review Corpus (Council et al., 2010)
- Plus: Italian (Attardi et al., 2015) and Spanish (Stricker et al., 2015)
1. Mark-up scheme for negation: negation cues, scope and focus
2. Treatment of negation reinforcement in romance languages
3. Annotation of:
   - Fact-Ita Bank: news stories from a regional Italian newspaper
   - tweets

(http://hlt-nlp.fbk.eu/technologies/fact-ita-bank)
NEGATION IN ITALIAN
Negation in Italian can be expressed by:

- **Adverbs**: *non*/*not* (“I do not live in Valencia”)
- **Pronouns**: *nessuno*/*nobody* (“Nobody left the room”)
- **Determiners**: *nessun*/*no* (“I see no trees around”)
- **Prepositions**: *senza*/*without*
- **Verbs**: *rifiutare*/*to refuse*
- **Nouns**: *rifiuto*/*refusal*
- **Prefixes**: *in-, dis-*
Features of negation in Italian

Italian is a romance language

- Some word forms vary in negative sentences (Qualcuno vs. nessuno)
- Reinforcement of negation

Example

1. **Nessuno ti ha cercato mentre eri via**
   - Italian: Nobody looked for you while you were away
   - English: ‘Nobody looked for you while you were away’

2. **Non ti ha cercato nessuno mentre eri via**
   - Italian: Not you has looked for nobody while you were away
   - English: ‘Nobody looked for you while you were away’
NEGATION CUE, SCOPE AND FOCUS
Decisions for annotation

Negation cues

· Only adverbs, pronouns, determiners and prepositions annotated
· Multiword negation cues are annotated as a single unit

Example

(3) Era una donna **per niente** candida.
    Was a lady **for nothing** candid.
    ‘*She was a woman not at all candid.’

· Negative forms that do not express negation not annotated. **Non appena**/“as soon as”.
Reinforcement

- Reinforcement annotated as secondary negation cue
- Linked to the main negation cue

Example

(4) I treni *non* arrivano *mai* in tempo.
The trains not arrive never on time.
‘The trains never arrive on time.’

(5) I treni *mai* arrivano in tempo.
Decisions for annotation

Scope

- Clause level annotation (but recoverable information from same sentence)
- Subordinated clauses included in the scope

Example

Non si sa ancora con essatezza *quanti siano stati i morti*. It is not exactly known yet *how many were the dead*. 
Decisions for annotation

Focus

- Some focus markers
  - Word order
  - Presence of the subject
  - Prosody (but we are working on written data)

Example

Lui non si è recato ai funerali.
He did not attend the funerals.
Problem 1: negation cues not included in the scope (Morante et al., 2011) but some negation cues can coincide with the extent of the focus.

Example

Nessuno ha cercato di fermare l’uomo.  
Nobody tried to stop the man.

Solution: negation cues which not only express negation but carry a richer semantic meaning are included in the scope.

Example

Nessuno ha cercato di fermare l’uomo.  
Nobody tried to stop the man.
Problem 2: in case of cleft sentences, the focus can be detached from the clause where the negation is placed.

Example

È dal 30 agosto che non si può più comprare.
It is since the 30th of August that it is no longer possible to buy it.

Solution: the extent of the scope is expanded to include as well the detached part of the cleft sentence.

Example

È dal 30 agosto che non si può più comprare.
It is since the 30th of August that it is no longer possible to buy it.
ANNOTATING NEGATION IN TEXTS
**Annotation of Fact-Ita Bank**

Data:

- Fact-Ita Bank is a news corpus in Italian
- 71 documents have been annotated
- Corpus already annotated with events and event factuality
  (http://hlt-nlp.fbk.eu/technologies/fact-ita-bank)

Annotation task:

1. Inter-annotator agreement (IAA): 12 documents
2. Annotation of 59 documents by 3 annotators
### Inter-annotator Agreement

<table>
<thead>
<tr>
<th></th>
<th>IAA - 1</th>
<th>IAA - 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>documents</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td># negation cues</td>
<td>47</td>
<td>30</td>
</tr>
<tr>
<td>negation cue</td>
<td>0.93</td>
<td>0.98</td>
</tr>
<tr>
<td>scope</td>
<td>0.52</td>
<td>0.67</td>
</tr>
<tr>
<td>focus</td>
<td>0.55</td>
<td>0.58</td>
</tr>
</tbody>
</table>

**Table:** IAA in terms of average pairwise F-measure.
Annotation of negation in tweets

Data:

- 301 tweets
- corpus already annotated with events and event factuality (http://facta-evalita2016.fbk.eu/pilot-task)

Example

Io non sono d’accordo che abbiano nominato grand Budapest hotel e il libro della vita.
I don’t agree with the fact that they have nominated grand Budapest hotel and The Book of Life.
## Quantitative Data

<table>
<thead>
<tr>
<th></th>
<th>news articles</th>
<th>tweet corpus</th>
</tr>
</thead>
<tbody>
<tr>
<td>docs</td>
<td>71</td>
<td>301</td>
</tr>
<tr>
<td>tokens</td>
<td>31,596</td>
<td>4,920</td>
</tr>
<tr>
<td>sentences</td>
<td>1,290</td>
<td>301</td>
</tr>
<tr>
<td>tweets/sent. w. neg.</td>
<td>278</td>
<td>59</td>
</tr>
<tr>
<td>negation cues</td>
<td>282</td>
<td>71</td>
</tr>
<tr>
<td>reinforcement</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>average size scopes</td>
<td>9.11</td>
<td>4.69</td>
</tr>
<tr>
<td>average size focus</td>
<td>3.2</td>
<td>1.61</td>
</tr>
</tbody>
</table>

**Table:** Quantitative data about the annotated corpora
DISCUSSION
Focus is the part of the sentence most prominently negated. But which is that part?
Example

(6) Un dolore che la notizia della cattura del colpevole non attenua

‘A pain that the news of the arrest of the culprit does not lessen’

Focus: “dolore/che”? “la notizia della cattura del colpevole”? “attenua”?
Example

(7) Avete vinto? No.
Did win? No.
‘Did you win? No.’

Focus and scope are part of a different sentence than the negation cue.

Avete vinto? No.


