The Particular Negative

A Distributional Study on Some Aspects of Meaning Contradicting Logical Equivalence

Alessandro Favaro

Università degli Studi di Padova

The Particular Negative

A Distributional Study on Some Aspects of Meaning Contradicting Logical Equivalence

Introduction

- What are particular negative statements?
- What types of particular negative statements did we take into account?

The Distributional Study

- How did we carry out the distributional study?
- Which aspects did we take into account?
- Results and final observations
- Possible future work

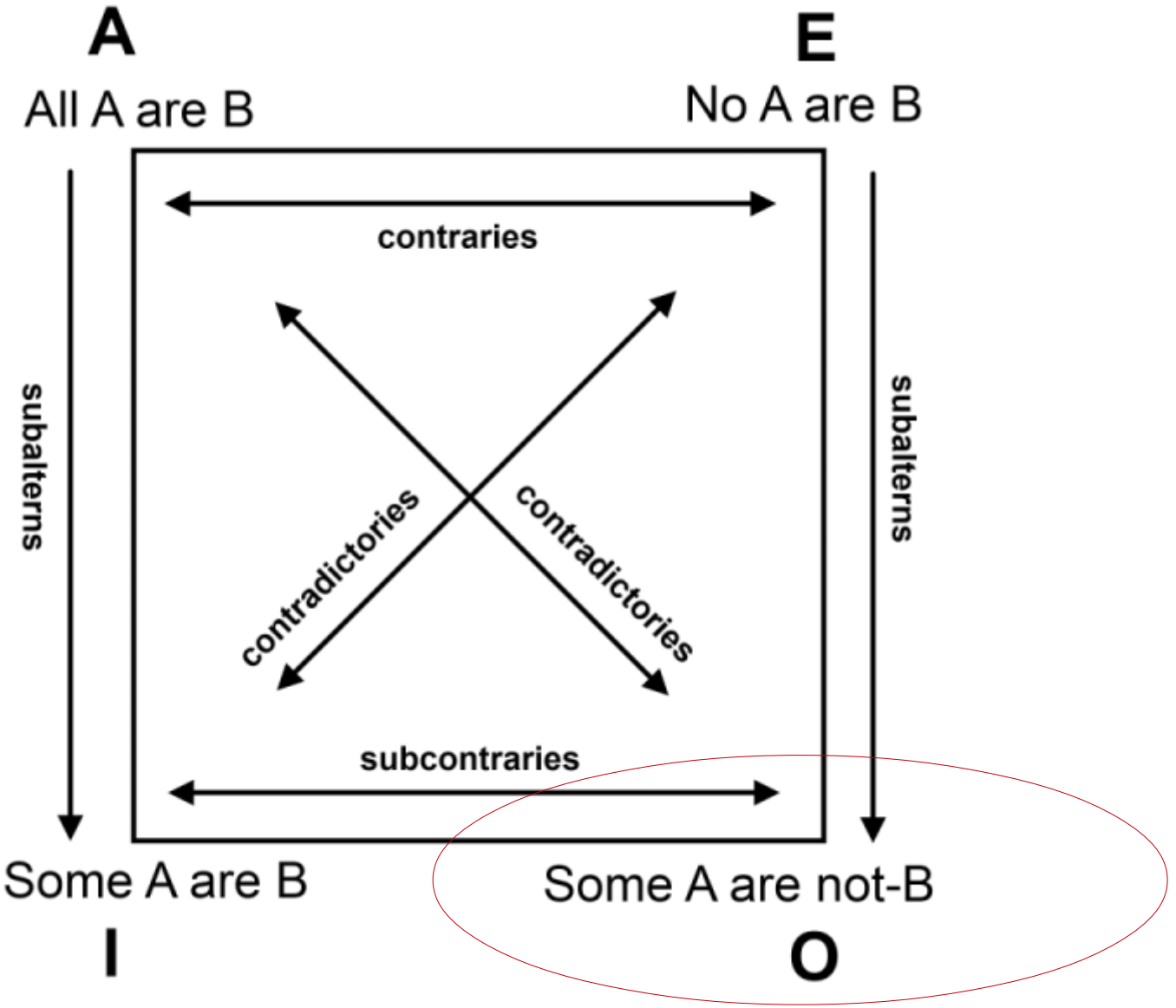
The Particular Negative

A Distributional Study on Some Aspects of Meaning Contradicting Logical Equivalence

Introduction

- What are particular negative statements?

The Aristotelian Square of Opposition



The Particular Negative

A Distributional Study on Some Aspects of Meaning Contradicting Logical Equivalence

Introduction

- What types of particular negative statements did we take into account?

The Particular Negative:

Focus on the **O corner** of the Aristotelian Square of Opposition

- We want to account for two types of particular negative expressions, i.e. two expressions involving quantifiers which are used within particular negative statements:
- *Not all;*
- *Some* followed by a verbal negation (*some not*).

Two example statements fo the two expressions under discussion are:

- *Not all birds can fly*;
- *Some birds cannot fly*.

Not all *birds can fly* vs. ***Some*** *birds can****not*** *fly*

Even though the two statements are logically equivalent, we want to check whether *not all* and *some not* are used in **different contexts** and for **different purposes**.

So, we want to look for those **aspects of meaning** related to the two expressions which might **contradict their logical equivalence**.

So, there might be some sort of incompatibility between two levels of meaning:

- *Logical*
- *Conversational*

logical level vs. *conversational level*

This possible incompatibility has already been highlighted by Jespersen(1924), Grice (1975), Horn (1989) within the treatment of the inference from *some* to *not all:*

Some birds can fly = Not all birds can fly

The inference from the first sentence to the second is widely recognised as an example of **scalar implicature** (Grice, 1975).

'[…] no special logical treatment of the inference from *some* to *not all* (*some not*) is required – which is just as well, since the context dependance and epistemic qualification associated with the inference would vitiate a logical treatment in any case.' (Horn, 1989: 213)

Hence, the theoretical assumption at the basis of the study:

On the one hand, two quantified expressions that are not equivalent from a logical point of view might be equivalent from a conversational one...

Some birds can fly = Not all birds can fly

...on the other hand, two quantified expressions that are equivalent logically might not be equivalent conversationally.

Not all birds can fly ≠ Some birds cannot fly

Where are we going to look for such differences and such aspects of meaning?

↓

Within the **different contexts of occurrence** of *not all* and *some not*

↓

We want to check these aspects by carrying out a **corpus-based distributional study** on *not all* and *some not*:

- In which kind of contexts do they usually occur?
- What are the differences between the contexts surrounding *not all* and *some not* ?

Distributional Semantics

Distributional Hypothesis:

Words and phrases which appear in **similar contexts** have **similar meanings**; in statistical terms, if an expression is likely to appear in a certain context, that context has an i**nfluence** on its meaning.

The theoretical basis of distributional semantics traces back to structuralist linguistics and British corpus linguistics (Harris, 1954; Firth, 1957).

Up-to-date accounts of the framework and methods of analysis related to distributional semantics can be found in Widdows (2004), Padò and Lapata (2007), Lenci (2008), Baroni and Lenci (2010).

The Particular Negative

A Distributional Study on Some Aspects of Meaning Contradicting Logical Equivalence

The Distributional Study

- How did we carry out the distributional study?

Choosing a suitable corpus...

The distributional study has been carried out on a written corpus; since we needed a corpus which is likely to capture at least some **conversational** aspects of language, we chose an **Intenet-based corpus.**

↓

Nowadays, it is widely recognised that texts on the web are mostly characterized by a **conversational style**, and various manuals of web writing (Felder, 2012; Fortis, 2013) recommend using language of this kind, as it is more suitable for web users.

...the ukWaC corpus

We explored the **ukWaC corpus**, which was prepared by Adriano Ferraresi; the construction was carried out by crawling the .uk Internet domain. It contains more than **2 billions tokens** and includes 'both **pre-Web texts**' and '**texts representing Web-based genres**, like personal pages, blogs and postings in forums' (Ferraresi et al., 2008).

Retrieving the data

The ukWaC corpus has been explored within the **CIMeC** (University of Trento) server, where the corpus itself is stored and subdivided into 24 parts; the **extraction** of the occurrences of *not all* and *some not* has been made by using the basic commands of **CQP (corpus query processor) language.**

Portion of corpus explored: first 12 parts (half corpus)

Portion of context explored: 5 sentences, i.e. the one including either *not all* or *some not*, the two preceding and the two following.

> set context 3s

Not all can vs. ***Some cannot***

Analysing every single ocurrence of *not all* and *some not*, even considering only a half of the corpus, would have been particularly difficult, as we would have had to deal with an enormous number of:

- Verb and noun phrases;
- Lexical and syntactic variables.

We decided to narrow the field of the research by formulating a **more specific command**:

> “not” %c “all” [] {1,2} “can”

> “some” %c [] {1,2} “cannot|can.t”

So, we specifically focused on two strings of the type ***not all can*** and ***some cannot***.

Cleaning the data

Not all the occurrences extracted from the ukWaC corpus matched the patterns we wanted to account for.

e.g.

- We have some more information about some of these companies in our shop, just pop in. But that's *not all* you *can* do!
- We just keep up: but at the expense of *some* tasks we *cannot* tackle.

So, we carried out a **'cleaning' process** by examining each single context of occurrence:

Total number of occurrences	**Before** the cleaning process	**After** the cleaning process
Not all can	493	**454**
Some cannot	623	**469**

Once the data were **retrieved** and **cleaned**, our analysis on the **923** contexts of occurrence of *not all can* and *some cannot* could finally get started!

The Particular Negative

A Distributional Study on Some Aspects of Meaning Contradicting Logical Equivalence

The Distributional Study

- Which aspects did we take into account?

1 – Context analysis (1)

Focus on the following aspects:

1) The **quantity** with which the **two expressions** and their **alternatives** (i.e. 'those who cannot' and 'those who can') are most likely to be associated;

2) Which one of the two expressions is most likely to be **specified** by means of **examples** within its context of occurrence;

1 – Context analysis (2)

3)Which one is more frequently used within **disjunctive structures** (involving words such as *but*, *however*, *although*...);

4)Which one is more likely to be based on **presuppositions** related to the same predicate as applied to ***all***. (For example, for statements such as *not all humans are honest* and *some humans are not honest* to be understood properly, they need to be related to the presupposition that *all humans are honest* represents an ideal situation.)

...presuppositions related to *all*

By using either *not all can* or *some cannot*, what is one likely to presuppose on the same predicate as applied to the whole quantity? We considered **three kinds** of presupposition:

- *All* is the **ideal situation.** It is signalled by adverbs such as *unluckily* and *unfortunately*, or expressions such as *it's a pity that*.
- *All* represents an **unachievable condition**. This reading can be triggered by adverbs such as *obviously* and *of course*.
- *All* is connected with a **possible wrong belief** of the addressee (introduced by words and phrases such as *note that, beware tha*t, *be careful*:,etc.).

2- Word co-occurrence analysis

What are the words occurring most frequently along with *not all can* and *some cannot* within the first half of the ukWaC corpus?

- For the purpose, we considered only the **single sentences** in which the two expressions are included.
- **Grammar words** (*the*, *that*, *if*, *of*, etc.) have **not** been taken into account.
- **Total number** of co-occurring words (wordcounttools.com):
  - 8576 for *not all can*;
  - 10500 for *some cannot*.

The Particular Negative

A Distributional Study on Some Aspects of Meaning Contradicting Logical Equivalence

The Distributional Study

- Results and final observations

Results – Context analysis

1) *Not all can* is followed by a **reference to its complement** (i.e. to those who/which can) more frequently than is *some cannot* (12,4% of the contexts for *not all can* and 6% for *some cannot*); moreover, such alternatives are often associated with quantifiers such as *most*, *many*, *the majority of*;

2) *Some cannot* is more frequently followed by **examples** specifying who or what *some* refers to (14,9% vs. 6,3%);

3) Both *not all can* and *some cannot* are very often used within **disjunctive structures** (34,1% and 21,3% of the contexts);

4) Presuppositions related to *all* underlie more frequently the use of *not all can* than that of *some cannot* (8,1%, 9,4%, 6,8% vs. 2,5%, 2,3%, 4,3%).

Not all can	Items referred to *not all*		Items referred to the alternative		Disjunctive structures	Presuppositions		
	Ex.	Q	Ex.	Q		1	2	3
	29 (6,3%)	22 (4,8%)	23 (5%)	34 (7,4%)	155 (34,1%)	37 (8,1%)	43 (9,4%)	31 (6,8%)
Some cannot	Items referred to *some*		Items referred to the alternative		Disjunctive structures	Presuppositions		
	Ex.	Q	Ex.	Q		1	2	3
	70 (14,9%)	17 (3,6%)	0	28 (6%)	100 (21,3%)	12 (2,5%)	11 (2,3%)	20 (4,3%)

Results – Word co-occurrence analysis (1)

Not all can		***Some cannot***	
Words	Occurrences	Words	Occurrences
will	41 (5,6%)	people	81 (8,5%)
people	28 (3,8%)	will	40 (4,2%)
time	23 (3,2%)	because	38 (4,0%)
afford	21 (2,9%)	things	28 (2,9%)
work	20 (2,7%)	need	26 (2,7%)
used	18 (2,5%)	afford	23 (2,4%)
same	17 (2,3%)	children	21 (2,2%)
us	17 (2,3%)	time	20 (2,1%)
due	16 (2,2%)	simply	19 (2,0%)
should	16 (2,2%)	others	19 (2,0%)
browsers	14 (1,9%)	always	18 (1,9%)

Results – Word co-occurrence analysis (2)

Not all can		***Some cannot***	
Words	Occurrences	Words	Occurrences
support	14 (1,9%)	still	18 (1,9%)
because	14 (1,9%)	without	17 (1,8%)
get	14 (1,9%)	being	17 (1,8%)
students	14 (1,9%)	cannot	17 (1,8%)
need	14 (1,9%)	work	16 (1,7%)
way	13 (1,8%)	just	16 (1,7%)
course	12 (1,6%)	get	16 (1,7%)
new	12 (1,6%)	well	14 (1,5%)
see	12 (1,6%)	like	14 (1,5%)
up	12 (1,6%)	example	14 (1,5%)
users	12 (1,6%)	easily	14 (1,5%)

Final observations (1)

1) *Some* is often specified by one or more given **examples** – a situation in which *not all* is involved less freqently.

↓

'Uncertain' status of *some* in terms of quantity

↓

If within a zero-context environment we have no clues for associating this quantifier with a given quantity, the speaker may need to build a context specifying more details about what *some* refers to: the **examples** may help the addressee resolve the uncertain status.

Final observations (2)

2) *Some cannot* is clearly less focused on the definition of an **alternative** than is *not all can*.

↓

No examples referred to the **alternative** of *some cannot* – that is, to *those who can* – have been found throughout the first half of the ukWaC corpus.

Furthermore, even if the alternatives of both *some cannot* and *not all can* are often associated with a large quantity, there are **more alternatives** of *not all can* which are connected with quantifiers such as *most*, *many*, *the majority of*.

Final observations (3)

3) Both *not all can* and *some cannot* are frequently involved in **disjunctive structures**.

↓

The **particular negative** is used quite often for either contrasting a given condition or introducing one that will then be contrasted.

4) All three **presupposition types** lie more frequently at the basis of the use of *not all can*: of course, *not all* focuses more clearly on the quantifier *all*.

↓

We may argue that some presuppositions play an important role in triggering the choice of the **focus of negation**.

The Particular Negative

A Distributional Study on Some Aspects of Meaning Contradicting Logical Equivalence

The Distributional Study

- Possible future work

Possible future work

- When the modal verb *can* is used within particular negative statements – as happens in all the cases I considered in the corpus-based study, is it more likely to be related to **epistemic** modality, to **deontic** modality or to an **ability**?
- With regard to the use of particular negative expressions within **disjunctive structures**, are such expression more likely to be found in the **main** clause or the **subjunctive** clause? In this respect, do *not all* and *some not* behave in different ways?
- Focusing on **languages other than English**, do particular negative statements occur in conversational contexts that are somehow 'similar' to those surrounding *not all* and *some not*?

Thank you!

...any questions??