Argumentation: Bayesian analysis and structured meanings

Grégoire Winterstein
Laboratoire de Linguistique Formelle – Université Paris 3 - Sorbonne Nouvelle

gregoire.winterstein@linguist.univ-paris-diderot.fr

BNLSP Workshop – ESSLLI 2013

What is an argument?

Most treatments of argumentation (e.g. in philosophy, AI, psychology or linguistics) agree on the following:

- An argument is an attempt to **persuade** an agent ...
- An argument targets a **conclusion** (a **goal**)
- An argument is potentially **defeasible**, i.e. arguments can:
  - be compared
  - undercut, refute, undermine each other
  - an argument has a given **strength** in favor of its conclusion

What is a good argument?

- **Classical view**: a good argument is **logically valid**
  - it is an acceptable form of deduction or induction
  - it avoids fallacies and non-valid reasoning
- **Practical view**: an argument is as good as it is **persuasive**.
- In Bayesian terms: a good argument **raises the degree of belief** in its conclusion.
- This can be achieved in any way, as long as it is effective.
  - Hahn & Oaksford (2007): fallacies such as the argument from ignorance or the *petitio principii* can prove quite convincing in the right situation.

Linguistic argumentation

- The Bayesian treatment of argumentation might appear rather trivial:
  - Everything is handled by the update mechanism, supposing that priors and joint probability distributions are known.
  - If the probability of some H is raised by asserting A, then A is an argument for H.
- But, cf. Anscombre & Ducrot (1983, p. 9):

  > *The argumentative possibilities in a discourse are tied to the global linguistic structure of the utterances and not just to the content they convey.*

- (1-a) and (1-b) have the same content, but (1-a) is a better argument for selling an insurance plan:

(1) a. Starting at only 29.9$ a month!
b. At least 29.9$ a month!

Outline

Today:

- Take a look at the way some linguistic items seemingly affect the update mechanism.
- **Hypothesis**: the structure of meanings plays a role in the way arguments are evaluated.
- Focus on clearly argumentative situations:
  - The debated subject (the goal) is clear.
  - The speaker is trying to give valuable information related to the goal.
- Therefore, no word on the process of abduction.

Argument from ignorance

(2) Ghosts exist because no one has proved that they do not.

- (2) is an **argument from ignorance** (*argumentum ad ignorantiam*), a classical fallacy.
- (2) is not very convincing and is thus a **bad argument**.
- **Classical explanation**: (2) is bad because it's a fallacy.
- However, this type of argument can be convincing:

  (3) Drug X is safe because no test has proved that it has undesirable side-effects.

- What is the difference between (3) and (2)?
  - Hahn & Oaksford (2007): the difference between good and bad arguments is not a question of structure but of **content**.

Properties of the argument from ignorance

- H&O's Bayesian approach (correctly) predicts that:
  1. Negative arguments can be convincing in the right context.

     (4) Drug *X* is not toxic: a recent study showed no undesirable side-effect.

  2. Generally, negative arguments are less convincing than their positive counterparts: (5) is better than (4) (for the opposite conclusion)

     (5) Drug *X* is toxic: a recent study showed undesirable side-effects.

- What happens if both type of arguments are presented in parallel?

$\Rightarrow$ Test with **only**.

The interpretation of *only*

(6) Lemmy only drinks Jack Daniels.

a. ⇝Lemmy drinks nothing apart from JD.

b. ⇝Lemmy drinks JD.

- The content (6-a) is analyzed as the main content of (6)
- The content (6-b) is analyzed in different ways:
  - As a main content (Atlas, 1993)
  - As a presupposition (Horn, 1972; Rooth, 1992; Klinedinst, 2005; Singh, 2008; Beaver & Clark, 2008; Beyssade, 2010)
  - As a scalar implicature (van Rooij & Schulz, 2004)
  - Both as a standard and a weak presupposition (Zeevat, 2011)
- I assume that (6-b) is a non-main content.

(7) Does Lemmy only drink Jack Daniels?

a. ⇝Lemmy drinks JD.

The argumentative dimension of *only*

- Ducrot (1973, pp. 272–273): French **seulement** (=**only**) is an argumentative operator, it marks an inversion of the orientation of its prejacent.

(8) a. Lemmy has a master's degree. $\underset{arg}{\rightsquigarrow}$ Hire him.

b. Lemmy only has a master's degree. $\underset{arg}{\rightsquigarrow}$ Do not hire him.

$\Rightarrow$ Consequences on arguments from ignorance?

Only and the argument from ignorance

(9) a. Three of the existing studies have found dangerous side effects associated with drug X.

b. $\underset{arg}{\rightsquigarrow}$ *Drug X is dangerous.*

- Add **only**:

(10) Only three of the existing studies have found dangerous side effects associated with drug X.

Non main content : three studies found dangerous side effects
Main content : the other studies found no side effects
$\Rightarrow$ **Argument from ignorance**

- Who wins?

Predictions

- Anscombre & Ducrot: (10) argues against its prejacent i.e. in favor of *drug X is not dangerous*.
- Hahn and Oaksford (H&O) postulate that the argumentative effects come out of a Bayesian update made with the total meaning conveyed by the utterance.
- Contents are not distinguished according to their nature (e.g. main or non main content)
- Prediction on (10):
  - Both contents are taken into account.
  - The positive argument of the prejacent "wins" because it's more convincing, i.e. (10) should argue like its prejacent.

Negative prejacent

- The prejacent itself can be a negative argument:

(11) Only three of the existing studies failed to find undesirable side effects.

a. **Non main content**: 3 studies failed to find side effects.
b. **Main content**: other studies found side effects.

- The main content is a positive argument, H&O predict the same as Ducrot: (11) should argue in favor of *drug X is dangerous*.

Experimental protocol

- Test of the predictions with a protocol similar to that of H&O.
- Subjects were presented with a small paragraph:

(12) Barbara is wondering whether she should take digesterol for her stomach troubles. Her friend Sophia tells her that (**only**) seven of the existing medical studies (*have found/failed to prove*) that digesterol has undesirable side effects.

- Question:

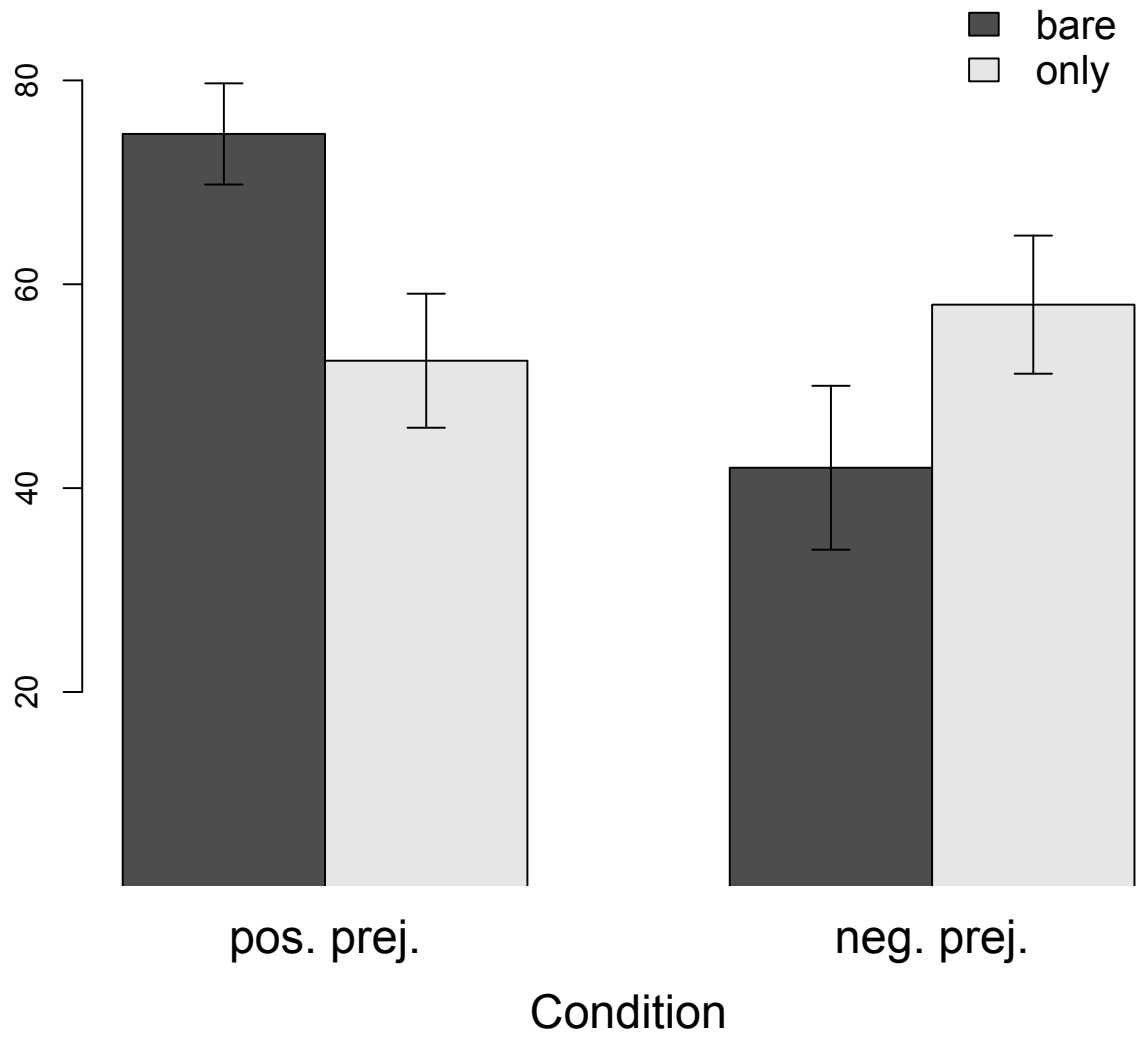
(13) How strongly do you think that Barbara is convinced that digesterol is dangerous given what Sophia tells her?

- Answer on a scale from 1 (*not convinced at all*) to 10 (*entirely convinced*)

Experimental protocol (cont.)

- Two binary factors
  1. Content of the prejacent: positive or negative argument.
  2. Absence/presence of **only**
- Questionnaire set up on `IbexFarm`.
- Eight target items.
- Ten fillers.
- Pseudo-randomization with a latin square design.
- Subjects recruited on the crowdsourcing platform `crowdflower`: 20 participants, paid 0.50$ for their participation.

Results



Results (cont.)

- The contribution of each predictor variable was assessed using model reduction and likelihood ratio test.
- Each factor has a significant contribution (presence of **only** $\chi^2(1) = 6.31, p = 0.01$; nature of the prejacent $\chi^2(1) = 8.22, p = 0.004$)
- Interaction between the two factors is significant ($\chi^2(1) = 9.52, p = 0.002$).
- This confirms that the use of **only** affects the argumentative properties of its host.

Only: taking stock

- Using **only** reverses the orientation of an argument.
- The reversal occurs in cases where the main content of the main argument is predicted to "lose" against the non-main content.
- One way of accounting for this is by postulating that only the main content of an utterance is taken into account when evaluating it.
- This has consequences on the abduction process: when the goal is unknown the search process is also guided by main content alone.

Interpretation of *almost*

Jayez & Tovena (2008):

(14) Lemmy is almost eighteen.

a. **Main content**: Lemmy is indistinguishably close to eighteen years old,

b. **Conventional Implicature**: Lemmy is not eighteen.

c. $18 - \varepsilon < age(Lemmy) < 18$

- **almost** negates its argument
- but keeps some of its argumentative properties:
  - (14) $\underset{arg}{\rightsquigarrow}$ Lemmy can drink alcohol.

Argumentative constraining

- Anscombre & Ducrot (1976); Jayez & Tovena (2008); Corblin (2012), **almost** blocks some continuations:

  (15) Few motorists go over 120 km/h, (# almost) 20%.

- This is expected if only the main content of "**almost** X" is taken into account when evaluating the argument:
  - Main content of **almost 20%** = a quantity Q s.t. $20\% - \varepsilon < Q$
  - This information alone cannot be used to argue that Q is small, hence the argumentative clash.

$\Rightarrow$ **Almost** constrains the argumentative orientation of its host.

Improvement effect of almost

- In some cases it appears that **almost** can **improve** the argument of its host: (16-b) > (16-a).

(16) a. A third of the persons who tested brand Y lost weight in the two weeks that followed.
b. Almost a third of the persons who tested brand Y lost weight in the two weeks that followed.

- The improvement seems to fade out if the host is a good argument:

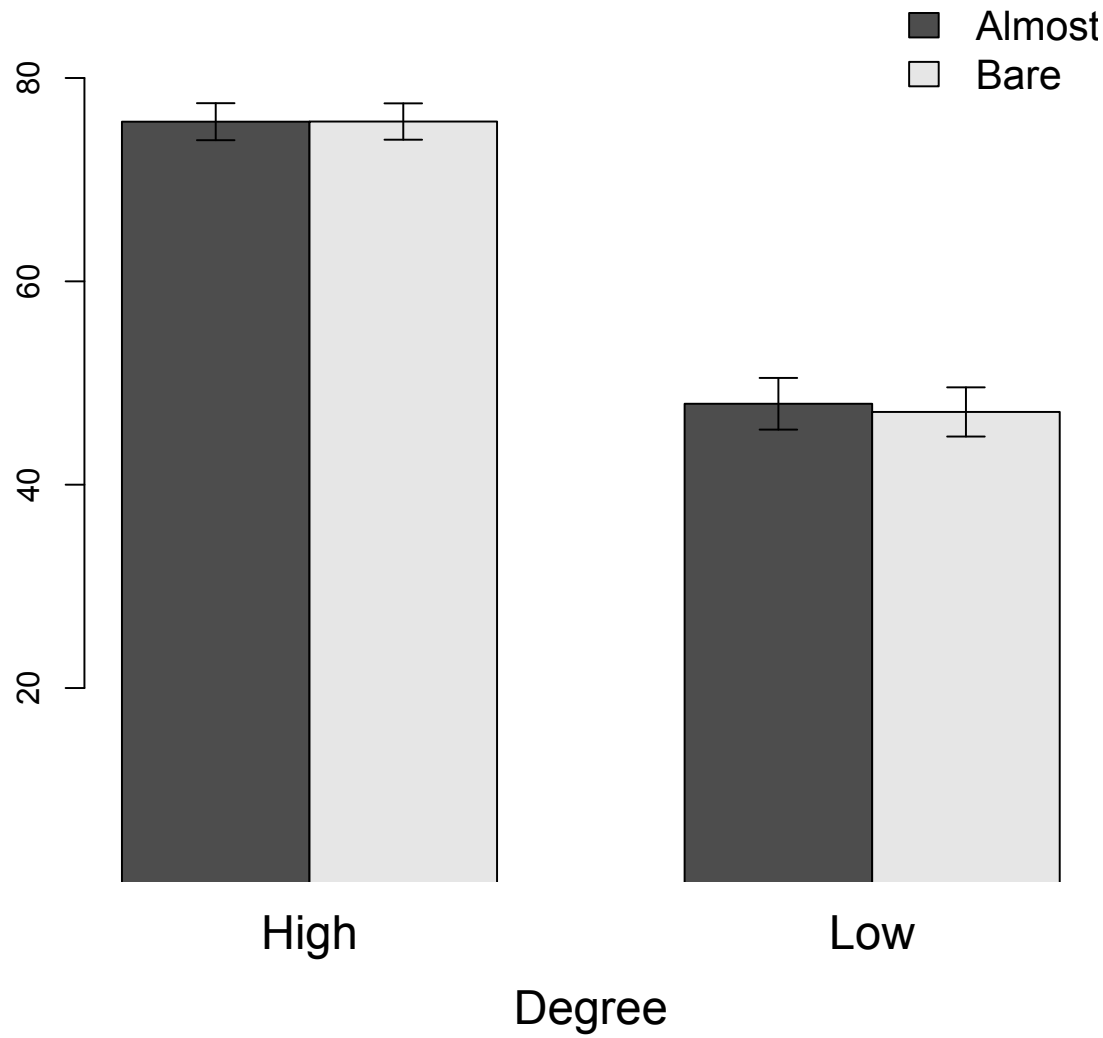
(17) a. 90% of the persons who tested brand Y lost weight in the two weeks that followed.
b. Almost 90% of the persons who tested brand Y lost weight in the two weeks that followed.

⇒ In (16-a) the argument is ambiguous between a positive and a negative one. Adding **almost** has a disambiguating effect.

Experiment

- Protocol similar to that of the previous experiment.
- Task: evaluate the degree of conviction given some information.
- 85 subjects recruited through `crowdflower`, all of them native speakers of American English.
- Two binary factors:
  - Presence/Absence of **almost**
  - High/Low proportion expressed.
- 20 item groups
- 40 fillers
- Actually, a series of experiments because few things worked the first time.

First Results



Results (cont.)

- Significant effect of the high/low proportion expressed ($\chi^2(1) = 48.21, p = 3.83e-12$)
- No effect of the use of **almost** ($\chi^2(1) = 1.187, p = 0.276$)
- No significant interaction between the factors.
- **Positive point**: in line with the hypothesis that **almost** does not degrade the properties of its host.
- **Negative point**: nothing very interesting here...

A closer look

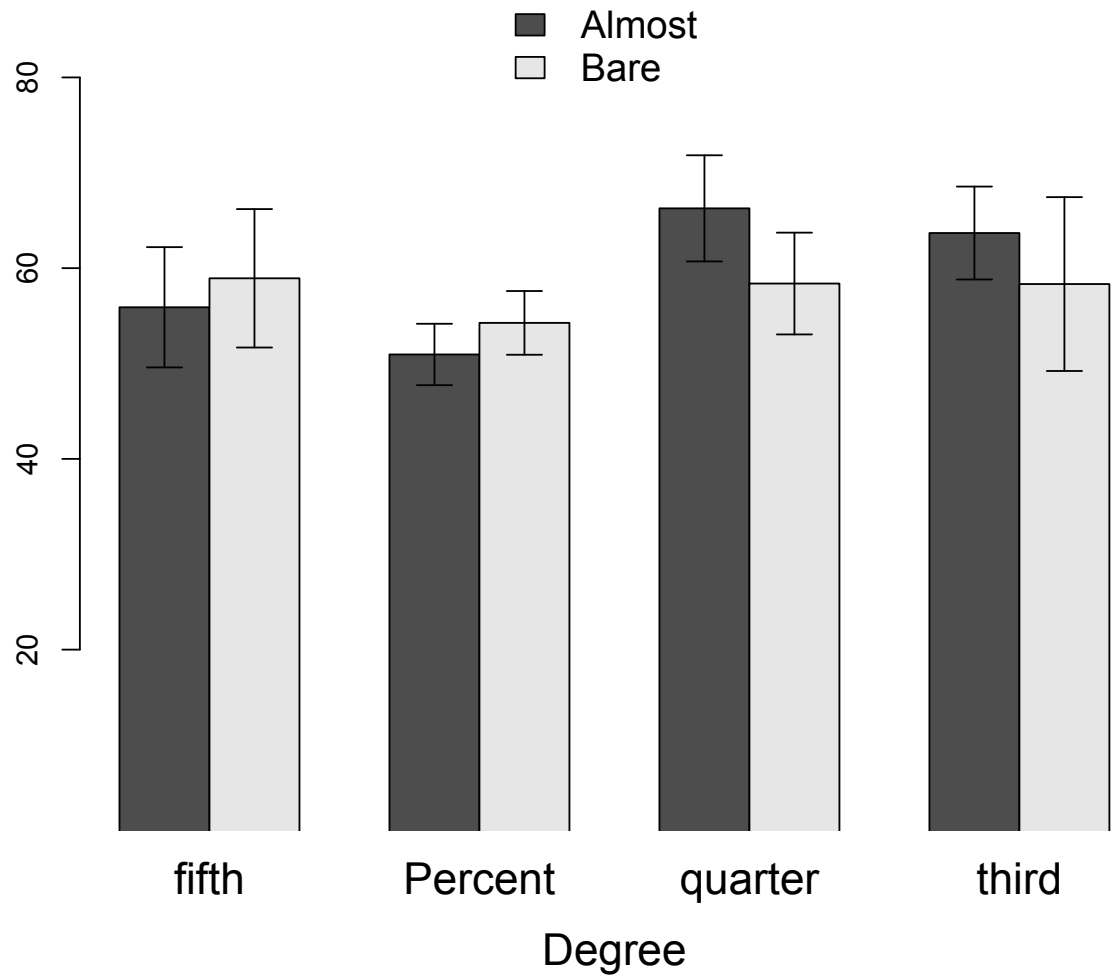
A post-hoc analysis shows several issues

1. Subjects develop strategies: restricting the analysis to the first half of the questionnaires for each subject yields significant results.
2. In order to vary the form of the target items, different ways were used to express proportions:

   Percentages *almost* **25%**
   Nouns *almost* **a fourth**

   The choice of expression appears to have an effect.
3. Some quantities were too low to be considered ambiguous, e.g. **a fifth** apparently created an argumentative clash when combined with **almost**.

Post-hoc: results by expression



Post-hoc analysis: results

- Focusing on the low degrees, the contributions of the following predictors were assessed using model reduction and likelihood ratio test:
  - Choice of expression: percentages, third, quarter, fifth
  - Presence/absence of **almost**
- **Results**:
  - No effect of the choice of expression alone ($\chi^2(1) = 4.48, p = 0.106$)
  - Significant effect of the use of **almost** ($\chi^2(1) = 27.65, p = 9.902e-07$)
  - Significant **interaction** between the two factors ($\chi^2(1) = 4.4737, p = 0.03442$)
- **Conclusion**: in some cases **almost** has the expected effect, i.e. it "improves" an argument even though the denoted quantity is inferior.
  - This is not true of all expressions, e.g. it does not work with percentages.
  - Even if the expressions appear equivalent when they appear without **almost**.

A theoretical issue

- **For all approaches**: explain the differences between the two classes of expression:

(18) a. almost 25% of the studies $<$ 25% of the studies
b. almost a quarter of the studies $>$ a quarter of the studies

- (18-b) is predicted by assuming that only the main content is taken into account.
- **Possible solution**: percentages are precise quantities unlike quarters and thirds.
- When **almost** modifies a precise quantity, its negative component is made more salient and is taken into account in the computation of argumentative effects, i.e. in *almost 25%*, the component *less than 25%* is more accessible than in *almost a quarter*.

Conclusion

- A Bayesian approach to argumentation is useful:
  - It allows for a uniform treatment of different types of argument.
  - Argumentation is part of a speaker's intention, and thus should be integrated in the process of interpretation of an utterance.
- The main content of an utterance plays a central role in the evaluation of the argument: it constrains the argumentative orientation of its host.
  - This allows "stronger" arguments to be ignored
  - This can improve arguments by disambiguation
  - This can lead to argumentative clashes
- However, non-main content might still have an effect (cf. percentages).This needs further elucidation.

Thank you

References I

Jean-Claude Anscombre, Oswald Ducrot (1976). "L'argumentation dans la langue". In: *Langages 42*, pp. 5–27.

— (1983). *L'argumentation dans la langue*. Liège, Bruxelles: Pierre Mardaga.

Jay David Atlas (1993). "The importance of being only". In: *Journal of Semantics 10*, pp. 301–318.

David I. Beaver, Brady Z. Clark (2008). *Sense and Sensitivity: How Focus determines meaning*. Wiley-Blackwell.

Philippe Besnard, Anthony Hunter (2008). *Elements of Argumentation*. Cambridge (MA): MIT Press.

Claire Beyssade (2010). "Seulement et ses usages scalaires". In: *Langue Française 165*, 1, pp. 103–124.

Francis Corblin (2012). "The meaning of *presque*". Presentation *Atelier DEI, Univ. Paris 4.*

Oswald Ducrot (1973). *La preuve et le dire*. Paris: Mame.

F. H. van Eemeren, R. Grootendorst, F. Snoeck Henkemans (1996). *Fundamentals of argumentation theory*. Mahwah, NJ: Erlbaum.

Ulrike Hahn, Mike Oaksford (2007). "The Rationality of Informal Argumentation: A Bayesian Approach to Reasoning Fallacies". In: *Psychological Review 114*, 3, pp. 704–732.

Larry Horn (1972). *On the Semantic Properties of Logical Operators in English*. Ph.D. thesis, Yale University.

Jacques Jayez, Lucia Tovena (2008). "Presque and almost: how argumentation derives from comparative meaning". In: Olivier Bonami, Patricia Cabredo Hofherr (eds.), *Empirical Issues in Syntax and Semantics*. CNRS, vol. 7, pp. 1–23.

Nathan Klinedinst (2005). *Scales and Only*. Master's thesis, UCLA.

Hugo Mercier, Dan Sperber (2011). "Why do humans reason? Arguments for an argumentative theory." In: *Behavioural and Brain Sciences 34*, pp. 57–111.

Arthur Merin (1999). "Information, Relevance and Social Decision-Making". In: L.S. Moss, J. Ginzburg, M. de Rijke (eds.), *Logic, Language, and computation*, Stanford: CSLI Publications, vol. 2, pp. 179–221.

References II

Henry Prakken (2010). "An abstract framework for argumentation with structured arguments". In: *Argument and Computation 1*, pp. 93–124.

Robert van Rooij, Kathrin Schulz (2004). "Only: meaning and implicatures". In: *Proceedings of Sinn und Bedeutung 9*. pp. 314–324.

Mats Rooth (1992). "A theory of focus interpretation". In: *Natural language semantics 1*, pp. 75–116.

Raj Singh (2008). *Modularity and Locality in Interpretation*. Ph.D. thesis, Massachusetts Institute of Technology.

Grégoire Winterstein (2010). *La dimension probabiliste des marqueurs de discours. Nouvelles perspectives sur l'argumentation dans la langue*. Ph.D. thesis, Université Paris Diderot.

Henk Zeevat (2011). "Expressing surprise by particles". manuscript. University of Amsterdam.

— (2013). "Accommodation in Communication". Manuscript, ILLC, University of Amsterdam.

— (In prep.). "Production and Interpretation of Natural Language". Book in preparation.