

Stratégies et opérateurs argumentatifs

Approche bayésienne (et expérimentale)

Grégoire Winterstein

Laboratoire de Linguistique Formelle – Université Paris 3
gregoire.winterstein@linguist.univ-paris-diderot.fr

LingLunch – 16 mai 2013

Plan

- Notion d'*argumentation* du point de vue psychologique :
 - Qu'est-ce qu'un bon argument ?
 - Comment modéliser l'efficacité d'un argument ?
- ⇒ Approche *bayésienne* : cas de l'argumentation par ignorance
- *Argumentation dans la langue* :
 - Relier les approches précédentes avec les approches linguistiques qualifiées d'argumentatives.
 - Test expérimental de l'hypothèse de "l'argumentation dans la langue" : cas de *only* et *almost*.

1 L'argumentation

Définir l'argumentation

Linguistique

- Anscombe & Ducrot (1983, p. 8)

Un locuteur fait une argumentation lorsqu'il présente un énoncé E_1 (ou un ensemble d'énoncés) comme destiné à en *faire admettre* un autre (ou un ensemble d'autres) E_2 .

Psychologie

- van Eemeren et al. (1996)

Argumentation is a verbal and social *activity* of reason aimed at *increasing* (or decreasing) the *acceptability* of a controversial standpoint for a listener or reader, by putting forward a constellation of propositions intended to justify (or refute) the standpoint before a "rational judge".

- Mercier & Sperber (2011) :

An argument is a *persuasive message*.

Intelligence Artificielle

- Besnard & Hunter (2008, p. 2)

An argument is a *set of assumptions* (i.e., information from which conclusions can be drawn), together with a *conclusion* that can be obtained by one or more *reasoning steps* (i.e., steps of deduction).

- Prakken (2010) : A est un argument ssi
 - A est un axiome dans la base de connaissance \mathcal{K} du locuteur

- Ou, si A_1, \dots, A_n sont des arguments dans \mathcal{K} , alors A peut être :
 - L'application d'une règle monotone : $A_1, \dots, A_n \rightarrow \psi$
 - L'application d'une règle non-monotone : $A_1, \dots, A_n \Rightarrow \psi$

Propriétés d'un argument

Un argument

- vise à persuader un allocataire/agent/ ...
- vise une conclusion, un *but*
- Est potentiellement *défaisable*, c.à.d. que des arguments peuvent :
 - être comparés
 - se supplanter par différents mécanismes (réfutation, sape ...)
- ⇒ un argument a une *force* en faveur de sa conclusion

1.1 Sophismes : l'argument d'ignorance

Bien argumenter ?

- **Approche classique** : un bon argument est *valide*, il doit :
 - Respecter les "lois" de la logique : déduction, ou forme valide d'induction
 - Éviter les raisonnements *non-rigoureux* : induction généralisante, abduction ...
 - Éviter les sophismes (*fallacies*) qui sont des raisonnements jugés *faux*
 - appel à l'ignorance
 - argument *ad hominem*
 - argument d'autorité
 - *petitio principii* (e.g. argument circulaire)
 - Hamblin (1970) : difficile/impossible d'établir une liste satisfaisante.
 - Également l'approche en intelligence artificielle.
- **Approche empirique** : un bon argument est *convaincant*.
 - En termes bayésiens, un argument convaincant *augmente le degré de croyance* en sa conclusion.

Appel à l'ignorance

- (1) Les fantômes existent parce que personne n'a prouvé qu'ils n'existent pas.
- (1) est un exemple d'*appel à l'ignorance* (*argumentum ad ignorantiam*), un cas d'école de sophisme.
 - Forme générale de l'appel à l'ignorance (Walton, 1996) :

(2) a. Si A était vrai (faux), on le saurait/le présumerait/aurait pu le prouver.
 b. A n'est pas connu/présumé/prouvé.
 c. A est (probablement) faux (vrai).
 - (1) paraît peu convaincant : c'est un *mauvais argument*.
 - **Hypothèse classique** : (1) est mauvais parce que c'est un sophisme.

Convaincre par ignorance

- Il existe des appels à l'ignorance qui sont (très) convaincants :

(3) Aucun test n'a montré que l'EGF causait une prolifération cellulaire anormale.
<http://www.bioeffect.fr/questions-responses-1.html>
- Code civil, articles 1382, 1383, 1389
 - Pour obtenir une indemnisation, il faut prouver : la faute, le préjudice, le lien de causalité.
 - Le lien de causalité est parfois difficile à prouver :
 - Pylônes électriques et santé du bétail

- Affaire du Mediator
- Appel à l'ignorance : en l'absence d'éléments prouvant le lien de causalité, on suppose qu'il n'existe pas.
- Qu'est-ce qui différencie ces appels à l'ignorance (convaincants) de (1) ?

1.2 Approche bayésienne

- Hahn & Oaksford (2007) : adoption d'un cadre bayésien.
 - Probabilité = degré de croyance
- Ce qui différencie les bons des mauvais arguments n'est pas leur forme, mais leur *contenu*.
- Un bon argument augmente la probabilité de sa conclusion, peu importe sa forme.
- Examen de plusieurs sophismes pour montrer qu'ils peuvent être de bons arguments :
 - appel à l'ignorance
 - argument circulaire
 - argument de la “pente savonneuse/fatale”

Appel à l'ignorance : trois propriétés

1. Les arguments négatifs peuvent être convaincants, mais restent généralement moins convaincants que les arguments positifs (à conclusion opposée) : (4-a) est moins convaincant que (4-b).
 - (4) a. X n'est pas toxique : une étude récente n'a mis en avant aucun effet secondaire dangereux.
 - b. X est toxique : une étude récente a mis en avant de dangereux effets secondaires.
2. Plus il y a d'arguments négatifs, plus la conclusion est jugée acceptable : (5) est meilleur que (4-a).
 - (5) X n'est pas toxique : dix études récentes n'ont mis en avant aucun effet secondaire dangereux.
3. Les croyances préalables des agents influencent l'acceptabilité de l'argument, cf. (1).

Approche bayésienne : toxicité d'un médicament

- Notations :
 - e : observation d'un effet toxique
 - $\neg e$: pas d'observation d'un effet toxique
 - T : Le médicament X est toxique
 - $\neg T$: Le médicament X n'est pas toxique
- Argument positif, examen de $P(T|e)$
 - (6) X est toxique : une étude récente a mis en avant des dangereux effets secondaires.
- Appel à ignorance, argument négatif, examen de $P(\neg T|\neg e)$
 - (7) X n'est pas toxique : une étude récente n'a mis en avant aucun effet secondaire dangereux.

Sensibilité et Spécificité

Les tests utilisés pour détecter des effets toxiques sont décrits par deux paramètres :

Sensibilité : $P(e|T)$, plus la sensibilité est haute, plus on a de chances d'un test positif en présence d'un effet toxique.

Spécificité : $P(\neg e|\neg T)$, capture le taux de “faux positifs”. Plus la spécificité est haute, moins le test a de chances d'être positif lorsqu'il n'y a pas d'effet toxique.

Les tests médicaux sont “optimisés” en termes de sensibilité et spécificité.

Force des arguments

– Notations :

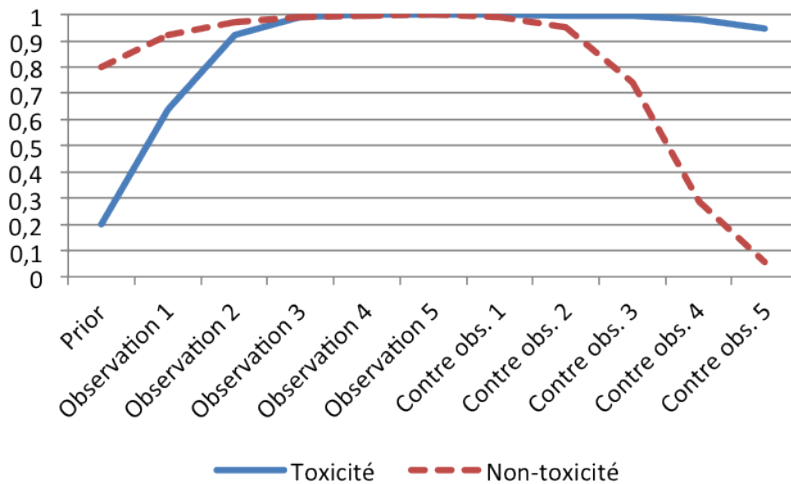
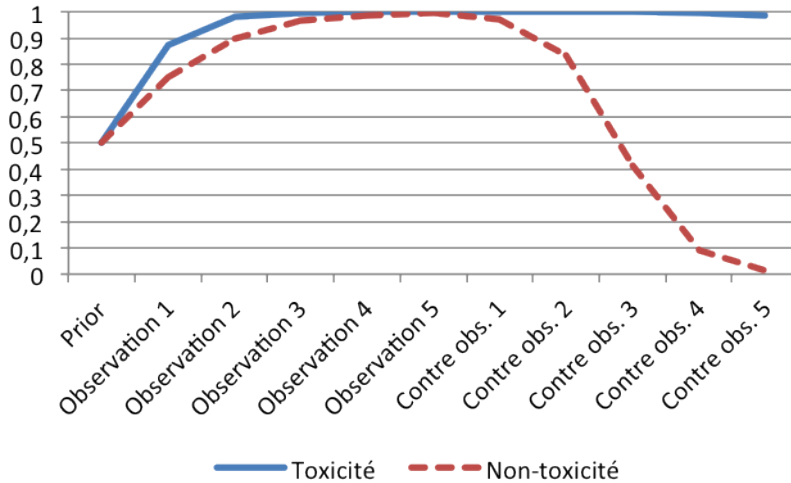
- n la sensibilité
- l la spécificité
- $h = P(T)$ le “prior”, c.à.d. la croyance préalable en la toxicité du médicament

– Par le théorème de Bayes :

$$(8) \quad \begin{aligned} \text{a.} \quad & P(T|e) = \frac{nh}{nh+(1-l)(1-h)} \\ \text{b.} \quad & P(\neg T|\neg e) = \frac{l(1-h)}{l(1-h)+(1-n)h} \end{aligned}$$

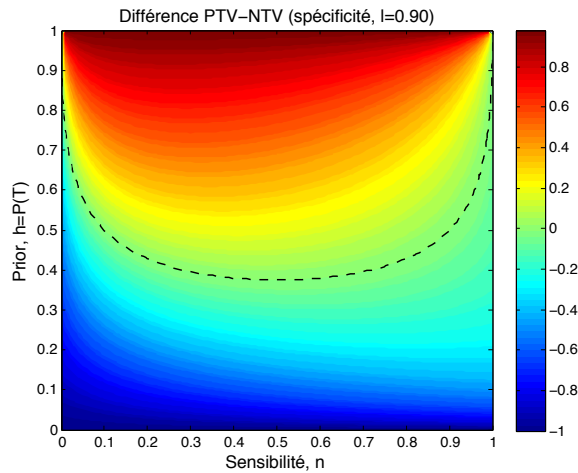
Illustration

Propriétés du test : *Spécificité* 0,9 *Sensibilité* 0,7



Généralisation

– Plus généralement, une observation positive (e) est un meilleur argument qu’une observation négative ($\neg e$) ssi $P(T|e) > P(\neg T|\neg e)$ (noté $PTV > NTV$), c.à.d. ssi : $h^2(n - n^2) > (1 - h)^2(l - l^2)$



Conclusion intermédiaire

Récapitulatif

- Une approche bayésienne permet de prévoir que :
 1. Les arguments négatifs peuvent avoir un réel pouvoir argumentatif.
 2. Généralement, un argument négatif est défait par un argument positif.
- Jusqu'ici on a supposé que les arguments étaient présentés en séquence.
- Que se passe-t-il lorsqu'on émet des contenus argumentativement contradictoires en même temps ?

2 Argumentation dans la langue

- Anscombe & Ducrot (1983, p. 9) :

Les enchaînements argumentatifs possibles dans un discours sont liés à la structure linguistique des énoncés et non aux seules informations qu'ils véhiculent.
- Les approches psychologiques ne tiennent (a priori) pas compte de la façon dont une information est véhiculée dans la prise en compte des effets argumentatifs.
- ⇒ Étude de cas mettant en jeu des informations argumentativement contradictoires.

2.1 Le cas de *seulement/only*

- (9) Lemmy boit seulement du Jack Daniels.
- a. \rightsquigarrow Lemmy ne boit rien d'autre que du JD.
 - b. \rightsquigarrow Lemmy boit du JD.
- Le contenu (9-a) est analysé comme le *contenu principal* de (9)
 - Le contenu (9-b) est analysé de différentes façons :
 - comme un contenu principal (Atlas, 1993)
 - comme une présupposition (Horn, 1972; Rooth, 1992; Krifka, 1999) ...
 - comme une implicature scalaire (van Rooij & Schulz, 2004)
 - À la fois une présupposition et une présupposition faible (Zeevat, 2011)
 - On le considère comme *contenu non-principal*.

La dimension argumentative de *seulement/only*

- Ducrot (1973, pp. 272–273) (pace (Winterstein, 2012)) : *seulement/only* est un opérateur argumentatif, il marque une *inversion* de l’orientation argumentative de son préjacent.

- (10) a. Le verre est à moitié plein. \rightsquigarrow Il faut le vider.
 b. Le verre est *seulement* à moitié plein. \rightsquigarrow Il faut le remplir.
 c. Cette voiture est chère, elle coûte 30000 euros.
 d. Cette voiture n’est pas chère, elle coûte *seulement* 30000 euros.

⇒ Quelles prédictions dans les cas mettant en jeu des appels à l’ignorance ?

Seulement et l’appel à l’ignorance

- (11) a. Trois des études existantes ont mis en avant des effets secondaires indésirables associés au médicament X.
 b. \rightsquigarrow *Le médicament X est dangereux.*

– Ajout de *seulement* :

- (12) Seulement trois des études existantes ont mis en avant des effets secondaires indésirables associés au médicament X.

Contenu non principal : trois études ont mis en avant des effets secondaires.

Contenu principal : les autres études existantes n’ont pas mis en avant d’effets secondaires (**Hypothèse** : le nombre de ces études est supérieur à trois, mais comparable.)

⇒ *Appel à l’ignorance*

Prédictions

- Prédiction de Anscombe & Ducrot : (12) doit argumenter contre son préjacent, c.à.d. en faveur de *Le médicament X est sans danger.*
- Hahn et Oaksford (H&O) supposent que les effets argumentatifs sont calculés en se basant sur la totalité du contenu véhiculé par un énoncé.
- A priori, ils ne distinguent pas l’information ajoutée en fonction de son statut informationnel.
- Prédiction sur (12) :
 - Les deux contenus doivent être pris en compte dans le calcul de la valeur argumentative de (12).
 - L’argument positif du préjacent “gagne”, c.à.d. que (12) doit argumenter dans le même sens que le préjacent.

Cas du préjacent négatif

– Cas où le préjacent est lui-même un argument négatif :

- (13) Seulement trois des études existantes n’ont trouvé aucun effet secondaire indésirable.
 a. *Contenu non-principal* : 3 études n’ont trouvé aucun effet secondaire.
 b. *Contenu principal* : des études (3 ou plus) ont trouvé des effets secondaires.

- Le contenu principal est un argument positif, l’approche de H&O rejoint celle de Ducrot : on prévoit que (13) argumente en faveur de *Le médicament X est dangereux.*
- Test des prédictions à partir d’un protocole emprunté à H&O.
- Présentation d’un texte (en anglais, validé par un locuteur natif) :

- (14) Barbara is wondering whether she should take digesterol for her stomach troubles. Her friend Sophia tells her that seven of the existing medical studies have found that digesterol has undesirable side effects.

– Question posée :

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

– Réponse sur une échelle de 1 (*not convinced at all*) à 10 (*entirely convinced*)

– Deux facteurs binaires :

1. Contenu du préjacent : argument positif ou négatif.

2. Absence/présence de *only*

– Mise en place du questionnaire sur **IbexFarm**.

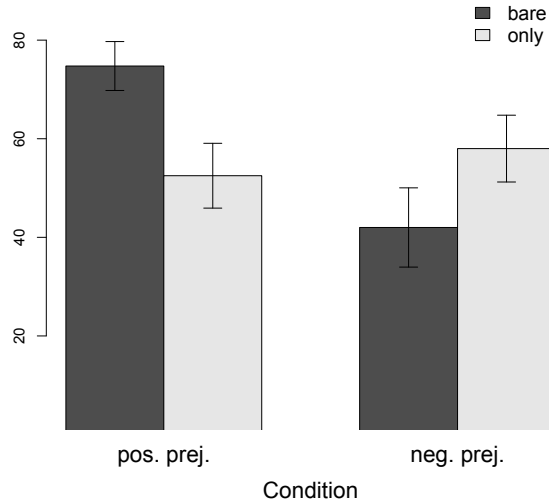
– Huit familles d'items cibles.

– Dix items distracteurs.

– Présentation pseudo-aléatoire des items en motif de carré latin.

– Recrutement de locuteurs via la plate-forme de crowdsourcing **crowdfower** : 20 participants, rémunérés 0,50\$ pour leur participation.

Résultats



Résultats (cont.)

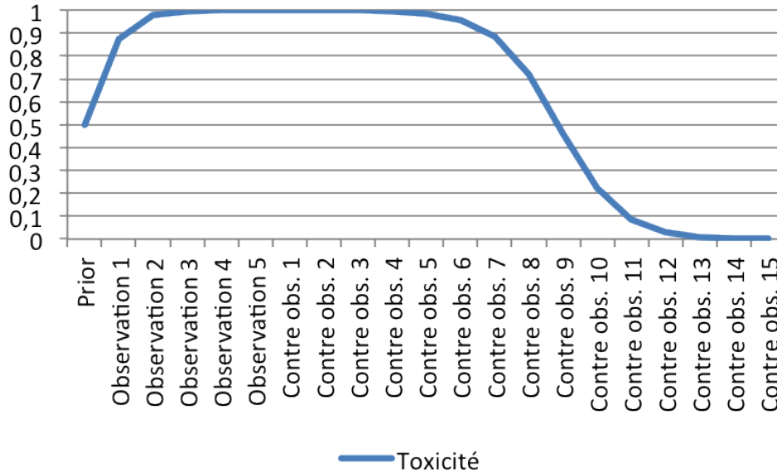
- La contribution de chacune des variables prédictrices a été mesurée par réduction de modèle et en utilisant le test de ratio de vraisemblance[©].
- Chaque facteur a une contribution significative (présence de *only* $\chi^2(1) = 6.31, p = 0.01$; nature du préjacent $\chi^2(1) = 8.22, p = 0.004$)
- L'interaction entre les deux facteurs s'avère significative ($\chi^2(1) = 9.52, p = 0.002$).
- Confirme que la présence de *only* affecte les propriétés argumentatives de son hôte.
- Dans une expérience pilote analogue menée en français (17 sujets), seule l'interaction est significative.

Explication alternative

- Les arguments positifs ne l'emportent sur les arguments négatifs que s'ils sont en nombres comparables.
- Il n'est pas évident de savoir quelle est la cardinalité de l'alternative considérée pour le préjacent de *seulement*.

- *seulement* trois des études ont montré X
 - Trois études ont montré X
 - Combien d'études n'ont pas montré X ?
- On peut faire l'hypothèse que *seulement* indique que cette cardinalité doit être suffisante pour permettre aux arguments négatifs de supplanter les positifs.

Combien d'arguments négatifs ?



2.2 Le cas de *presque/almost*

Jayez & Tovera (2008) :

- (16) Lemmy a presque dix-huit ans.
- Contenu principal* : Lemmy a un âge indiscernable (à gauche) de dix-huit ans.
 - Implicature conventionnelle* : Lemmy n'a pas dix-huit ans.

- *presque* permet de nier son argument
- et de garder les propriétés argumentatives de cet argument (pour les conclusions favorisées par des hauts degrés) :
 - (16) $\overset{arg}{\rightsquigarrow}$ Lemmy peut boire de l'alcool.
 - (16) $\overset{arg}{\not\rightarrow}$ Lemmy est jeune.

Effet mélioratif de *presque*

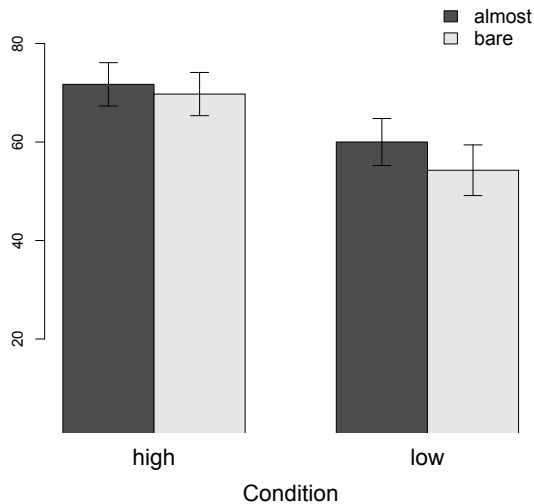
- En observant certains distracteurs de l'expérience précédente il apparaît que *presque* pourrait *améliorer* l'argument exprimé par son hôte : (17-b) paraît meilleur que (17-a) (cf. aussi J&T) :
 - Un tiers des personnes qui ont testé le produit Y ont perdu du poids dans les deux semaines qui ont suivi.
 - Presque un tiers des personnes qui ont testé le produit Y ont perdu du poids dans les deux semaines qui ont suivi.
- Intuitivement, l'effet mélioratif semble disparaître/être atténué lorsque l'argument exprimé par l'hôte est déjà bon.

- (18) a. 90% des personnes qui ont testé le produit Y ont perdu du poids dans les deux semaines qui ont suivi.
- b. Presque 90% des personnes qui ont testé le produit Y ont perdu du poids dans les deux semaines qui ont suivi.

Expérience

- Expérience pilote, protocole similaire au précédent.
- Tâche : évaluer le degré de conviction d'un individu étant donné une information qui lui est fournie.
- 40 sujets recrutés via **crowdfunder**
- Deux facteurs binaires :
 - Présence/Absence de *almost*
 - Haute/Basse proportion exprimée.
- 8 groupes d'items
- 16 distracteurs

Résultats



Résultats (cont.)

- Effet significatif de la proportion indiquée ($\chi^2(1) = 9.86, p = 0.002$)
- Effet marginal de l'emploi de *almost* ($\chi^2(1) = 2.84, p = 0.09$)
- Pas d'effet significatif de l'interaction des deux facteurs ($\chi^2(1) = 1.16, p = 0.28$).
- **Point positif** : cohérent avec l'hypothèse que *almost* n'affecte pas négativement les propriétés de son hôte.
- **Futur** : expérience plus solide en préparation (plus d'items).
- Si les résultats confirment que *almost* a un effet mélioratif :
 - Confirmation de certaines prédictions de J&T
 - À raffiner s'il s'avère que l'effet mélioratif est restreint aux proportions faibles.

3 Conclusion

Récapitulatif

- Intérêt de relier les approches bayésiennes de l’argumentation issues de la psychologie avec l’approche linguistique (encore plus en considérant les approches linguistiques également bayésiennes, cf. Merin (1999); Winterstein (2010)).
- Des expériences préliminaires semblent aller dans le sens de l’hypothèse selon laquelle le contenu argumentatif d’un énoncé n’est évalué qu’à l’aune de son contenu principal :
 - la contribution du préjacent de *seulement/only* est ignorée/minorée.
 - de même pour celle de l’implicature conventionnelle de *presque*.

Perspectives

- Quelles conséquences pour le processus d’interprétation d’un énoncé ?
 - La totalité de l’information véhiculée doit être ajoutée aux connaissances partagées.
 - La mise à jour des probabilités se fait par le biais d’une conditionalisation par le contenu principal de l’énoncé uniquement.
- Zeevat (In prep., 2013) : l’interprétation d’un énoncé est le fruit de la *reconnaissance d’une intention* (par un processus abduction).
- La reconnaissance de l’intention ne se fait qu’à partir du contenu principal d’un énoncé (?)

Bibliographie

- Jean-Claude ANSCOMBRE, Oswald DUCROT (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.
- Philippe BESNARD, Anthony HUNTER (2008). *Elements of Argumentation*. Cambridge (MA) : MIT Press.
- 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.
- Charles Leonard HAMBLIN (1970). *Fallacies*. London : Methuen.
- 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.
- Manfred KRIFKA (1999). “Additive particles under stress”. In : *Proceedings of SALT 8*. Cornell : CLC Publications, pp. 111–128.
- 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.
- 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.
- D. N. WALTON (1996). *Arguments from ignorance*. Philadelphia : Pennsylvania State University Press.
- Grégoire WINTERSTEIN (2010). *La dimension probabiliste des marqueurs de discours. Nouvelles perspectives sur l’argumentation dans la langue*. Ph.D. thesis, Université Paris Diderot.

— (2012). “‘Only’ without its scales”. In : *Sprache und Datenverarbeitung 35-36*, pp. 29–47.

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.