



Observer les discussions conflictuelles en ligne. Un cadre de pragmatique de corpus appliqué à Wikipédia

CÉLINE POUDAT

BCL, UNIVERSITÉ CÔTE D'AZUR

Plan

La situation communicative et le dispositif Wikipédia : les conflits à l'aune de la coopération

Wikipédia comme corpus: développement et explorations du corpus
Wikiconflits

Saisir et décrire les interactions conflictuelles dans Wikipédia: systèmes de marqueurs linguistiques du conflit dans Wikipédia

Perspectives

La situation communicative et le dispositif Wikipédia

LES CONFLITS À L'AUNE DE LA COOPÉRATION

Wikipédia

Une vingtaine d'années d'existence et un succès concluant (6^e site le plus visité en France – Alexa septembre 2021)

Un objet web 2.0 tentaculaire et hétérogène

- Multitude de pages relevant de différents genres du discours (articles encyclopédiques, pages d'aide, pages de discussion, etc.)
- Dimension génétique avec les historiques des révisions

Wikipédia

Un objet qui a été particulièrement observé par les sciences sociales (gouvernance, coopération, **conflits**, qualité; voir par exemple Viegas et al. 2004 et 2007, Brandes & Lerner 2007, Kittur & Kraut 2008 et 2010, Stvilia et al. 2008)

Un objet moins étudié par les sciences du langage

- Wikipédia comme corpus?
- Les articles sont plus décrits et exploités (notamment par le TAL) que les pages de discussion



Article page



WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Current events
Random article
About Wikipedia
Contact us
Donate

Contribute
Help
Learn to edit
Community portal
Recent changes
Upload file

Tools
What links here
Related changes
Special pages
Permanent link
Page information
Cite this page
Wikidata item

Print/export
Download as PDF
Printable version

In other projects
Wikimedia Commons

Languages 
Català
Deutsch
Español

Not logged in [Talk](#) [Contributions](#) [Create account](#) [Log in](#)

Article [Talk](#)

[Read](#) [Edit](#) [View history](#)



Pragmatics

From Wikipedia, the free encyclopedia

This article is about the subfield of linguistics. For the journal, see [Pragmatics \(journal\)](#). For the philosophy topic, see [Pragmatism](#).

In [linguistics](#) and related fields, **pragmatics** is the study of how [context](#) contributes to meaning. Pragmatics encompasses phenomena including [implicature](#), [speech acts](#), [relevance](#) and [conversation](#).^[1] Theories of pragmatics go hand-in-hand with theories of [semantics](#), which studies aspects of meaning which are [grammatically](#) or [lexically](#) encoded.^{[2][3]} The ability to understand another speaker's intended meaning is called *pragmatic competence*.^{[4][5][6]} Pragmatics emerged as its own subfield in the 1950s after the pioneering work of J.L. Austin and Paul Grice.^{[7][8]}

Contents [hide]

- Etymology
- Origins of the field
- Areas of interest
- Ambiguity
- Referential uses of language
 - Referential expressions in conversation
- Nonreferential uses of language
 - Silverstein's "Pure" Indexes
 - The performative
 - Jakobson's six functions of language
- Related fields
- Formalization
- In literary theory
- Significant works and concepts
- See also
- Notes
- References
- External links

Etymology [\[edit \]](#)

The word *pragmatics* derives via Latin *pragmaticus* from the Greek πραγματικός (*pragmatikós*), meaning amongst others "fit for action",^[9] which comes from πράγμα (*práγμα*) "deed, act" (in modern Greek πράγμα *práγμα* "an object, a thing that can be perceived by the senses"),^[10] in turn from the verb πράσσω (*prassō*, Attic Greek πράττω - *práttō*) "to do, to act, to pass over, to practise, to achieve."^[11]

Part of a series on
Linguistics

[Outline](#) · [History](#) · [Index](#)

General linguistics [\[show\]](#)

Applied linguistics [\[show\]](#)

Theoretical frameworks [\[show\]](#)

Topics [\[show\]](#)

 [Portal](#)

[V](#) · [T](#) · [E](#)

Talk page

Contents [hide]

- 1 Pragmatics subordinated
- 2 Intent or pragmatics with random, success, selection
- 3 Vague name references
- 4 Pragmatics != pragmatism
- 5 Moves and merges
- 6 Vandalism
- 7 Examples please!
- 8 linking Austin's acts to the Jakobson's uses of language
- 9 Can we make this article really good?
- 10 Messy ending
- 11 Definition of Pragmatics
- 12 Literal Meaning (Recanati, Searle)
- 13 Request
- 14 Philosophy Section
- 15 Minor huh in Origin
- 16 Tigers are obligate carnivores, not omnivores.
- 17 Citations?
- 18 Deictic Expressions
- 19 Citations
- 20 Misleading Information
- 21 Class Project

Pragmatics subordinated [edit]

Isn't it highly debateable to state, firmly in the beginning of the article, that pragmatics is a subfield of linguistics? --Xact (talk) 20:35, 19 December 2009 (UTC)

Why is that debatable? Why wouldn't pragmatics be considered a subfield of linguistics? KadedraDobyne (talk) 22:14, 3 February 2018 (UTC)

Speech pragmatics is also a field related to speech pathology. It is taught to autistic children. It's not just context. It has to do with nonverbal communication. <https://www.asha.org/public/speech/development/Pragmatics/> — Preceding unsigned comment added by Skysong263 (talk • contribs) 17:15, 5 October 2019 (UTC)

Intent or pragmatics with random, success, selection [edit]

Put a bunch of marbles in a bag. Pick one out without looking at it. Would this be a random selection or a selection at random ? It would be a selection at random but the intent is that consciousness is involved. Usually the symbol string "random" is not associated with consciousness. The symbol string *randomness* is usually associated with purposelessness.

La page de discussion Wikipédia: premiers éléments de description

Tous les articles ne nécessitent pas discussion (10% des articles de Wikipédia ont une page de discussion non vide)

Les contributeurs peuvent également commenter leurs modifications et donc interagir

La discussion a pour objectif de permettre aux contributeurs de se coordonner et de faire avancer l'article.

La discussion n'est pas un genre autonome: elle se réfère à l'article dans son état courant (dimension métadiscursive de discours sur le discours)

Exemple de commentaires dans l'historique des révisions (page

- (actu | diff)  5 août 2019 à 18:04 Erik Bovin (discuter | contributions) . . (385 590 octets) **(+741)** . . (→*Environnement* : précision : pas de retrait possible formellement avant le lendemain de l'élection présidentielle) (Balise : Éditeur de source 2017)
- (actu | diff)  25 juillet 2019 à 20:37 Cheep (discuter | contributions) . . (384 849 octets) **(+108)** . . (→*Sports*) (Balise : PHP7)
- (actu | diff)  25 juillet 2019 à 14:22 Lspiste (discuter | contributions) **m** . . (384 741 octets) **(-197)** . . (Annulation de la *modification* de *Sthubertliege* (d))l'article étant déjà beaucoup trop lourd à charger à cause de sa taille excessive, merci de ne pas ajouter de photo supplémentaire - surtout quand elles sont similaires à de nombreuses autres déjà présentes) (Balise : Annulation)
- (actu | diff)  25 juillet 2019 à 12:39 Sammyday (discuter | contributions) . . (384 938 octets) **(+21)** . . (Révocation des modifications de *Sthubertliege* (retour à la version précédente de *Sthubertliege*) : non conforme au reste de l'article - le RI le présente)
- (actu | diff)  25 juillet 2019 à 12:24 Sthubertliege (discuter | contributions) . . (384 917 octets) **(-21)** . . (Aucune de raison de mettre dans la description générale qu'il emploie des "contre-vérités". Sérieusement, il est très loin d'être le premier homme politique à mentir pendant une campagne électorale. Merci de ne pas faire de cette page un lieu anti-Trump et de rester objectif.) (Balise : Éditeur visuel)
- (actu | diff)  25 juillet 2019 à 12:11 Sthubertliege (discuter | contributions) . . (384 938 octets) **(+197)** . . (Balise : Éditeur visuel)
- (actu | diff)  25 juillet 2019 à 12:04 Sthubertliege (discuter | contributions) . . (384 741 octets) **(+63)** . . (Balise : Éditeur visuel)
- (actu | diff)  24 juillet 2019 à 13:54 Hercule (discuter | contributions) . . (384 678 octets) **(0)** . . (*typo*) (Balise : PHP7)
- (actu | diff)  11 juillet 2019 à 21:12 Guillaume Le Conquistador (discuter | contributions) **m** . . (384 678 octets) **(+5)** . . (→*Médiatisation et carrière dans les médias* : mise en forme)
- (actu | diff)  11 juillet 2019 à 21:04 Restefond (discuter | contributions) **m** . . (384 673 octets) **(-27)** . . (Annulation de la *modification* de *Restefond* (d) : oups trop tard) (Balise : Annulation)
- (actu | diff)  11 juillet 2019 à 21:02 Restefond (discuter | contributions) **m** . . (384 700 octets) **(+27)** . . (→*Positionnement des médias* : ajout d'un ref name pour corriger un problème dans la section suivante)
- (actu | diff)  11 juillet 2019 à 21:01 Cheep (discuter | contributions) **m** . . (384 673 octets) **(-51)** . . (*idem*, pas de source (« protectionniste » est assez proche)) (Balise : PHP7)
- (actu | diff)  11 juillet 2019 à 20:56 Hégésippe Cormier (discuter | contributions) **m** . . (384 724 octets) **(+176)** . . (→*Élection* : réparation d'un appel de référence : avec l'URL ce sera mieux...) (Balise : PHP7)

Discuter dans les règles

Les interactions sont des conduites ordonnées qui se déroulent selon certains schémas préétablis et obéissent à certaines règles de procédure

- Systèmes de droits et de devoirs, et donc d'attentes, susceptibles d'être satisfaites ou d'être contrariées
- La discussion est soumise à un ensemble de règles explicitement établies par Wikipédia

Les 5 piliers de Wikipédia

Wikipédia est une **encyclopédie**

Wikipédia recherche la **neutralité de point de vue**

Wikipédia est publiée sous **licence libre**

Wikipédia suit des **règles de savoir-vivre**

Wikipédia n'a pas d'autres règles que les cinq principes fondateurs énoncés ici (**droit à être audacieux, droit à l'erreur**)

Source:

https://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Principes_fondateurs

Les Normes Wikipédiennes concernant la tenue des discussions

EXEMPLES DE RECOMMANDATIONS

Discuter au lieu de révoquer

Être *amical, poli, positif*

Demander une source solide pour étayer une position est souvent bien plus efficace que d'argumenter contre cette position.

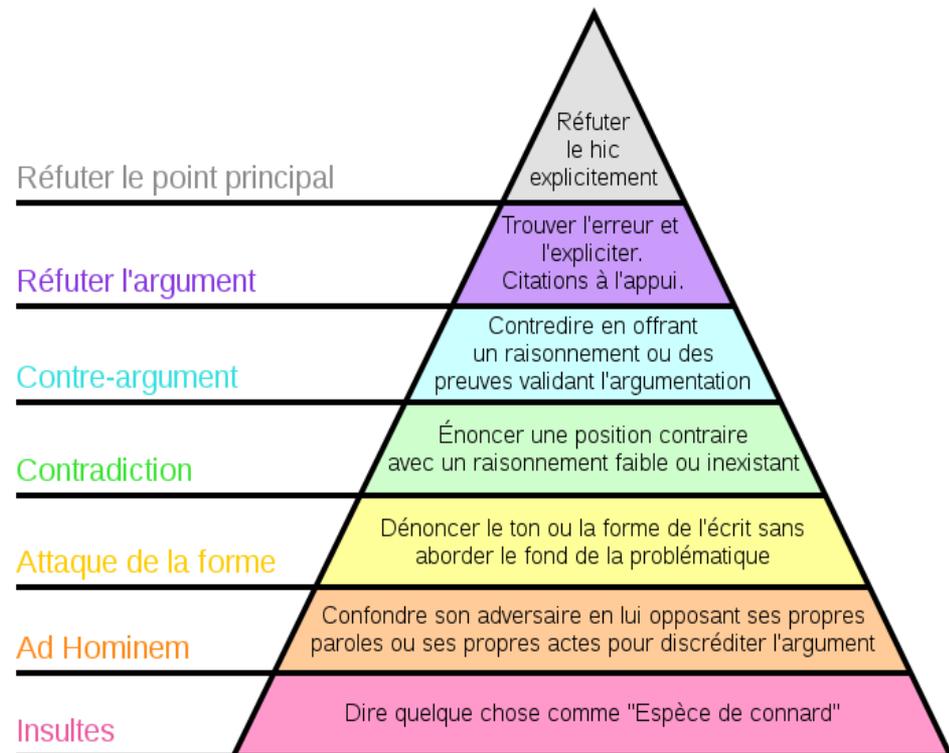
Demander de l'aide pour trouver des sources

Wikipédia n'est pas un forum

Source:

<https://fr.wikipedia.org/wiki/Aide:Discussion>

OBJECTIF: « RESTER DANS LE HAUT DE LA PYRAMIDE »



Impact des normes sur les Wikipédiens

Dispositif encourageant le développement d'une communauté virtuelle, cf. les neuf conditions proposées par Marcocchia (2016) : sentiment d'appartenance, ritualisation de la parole, stabilisation de rôles dans le groupe, familiarité, réciprocité, partage des valeurs et des finalités, histoire conversationnelle commune, persistance des échanges, constitution d'une charte de groupe.

Impact des normes sur les discussions

Rappels réguliers des normes et des principes de Wikipédia dans les pages de discussion

- Ce que Wikipédia n'est pas: « Wikipédia n'est pas un forum », « Wikipédia n'est pas une tribune », etc.
- Question des sources, problèmes de NPOV (neutralité de point de vue)

Développement de stratégies spécifiques d'expression indirecte du désaccord

- *Une source à ces affirmations?* (page Bogdanoff)
- *L'exposition du point de vue de Popper n'est pas neutre* (page Psychanalyse)

Impact du format Wiki sur la discussion

Particularités de la syntaxe wiki: les interactions s'éditent, il n'y a pas de bouton *send* comme dans les forums

- [Postings are] stretches of text that an individual user produces in private and then passes on to the server through performing a “posting” action (usually by hitting the [enter] key on the keyboard or by clicking on a [send] or [submit] button on the screen).
- Recommandations officielles de bonnes formation des discussions, mais on observe qu'elles ne sont pas systématiquement observées (du moins dans le Wikipédia francophone)
 - E.g. messages mal formés, sans signature, discontinus (édition d'un message précédent à différents endroits)

La conception de la discussion peut varier: liste de choses à faire, journal de bord...

Éléments de démographie et gouvernance

DEUX GRANDES CATÉGORIES DE PARTICIPANTS

les médiateurs, ou arbitres de la dispute (en vert)

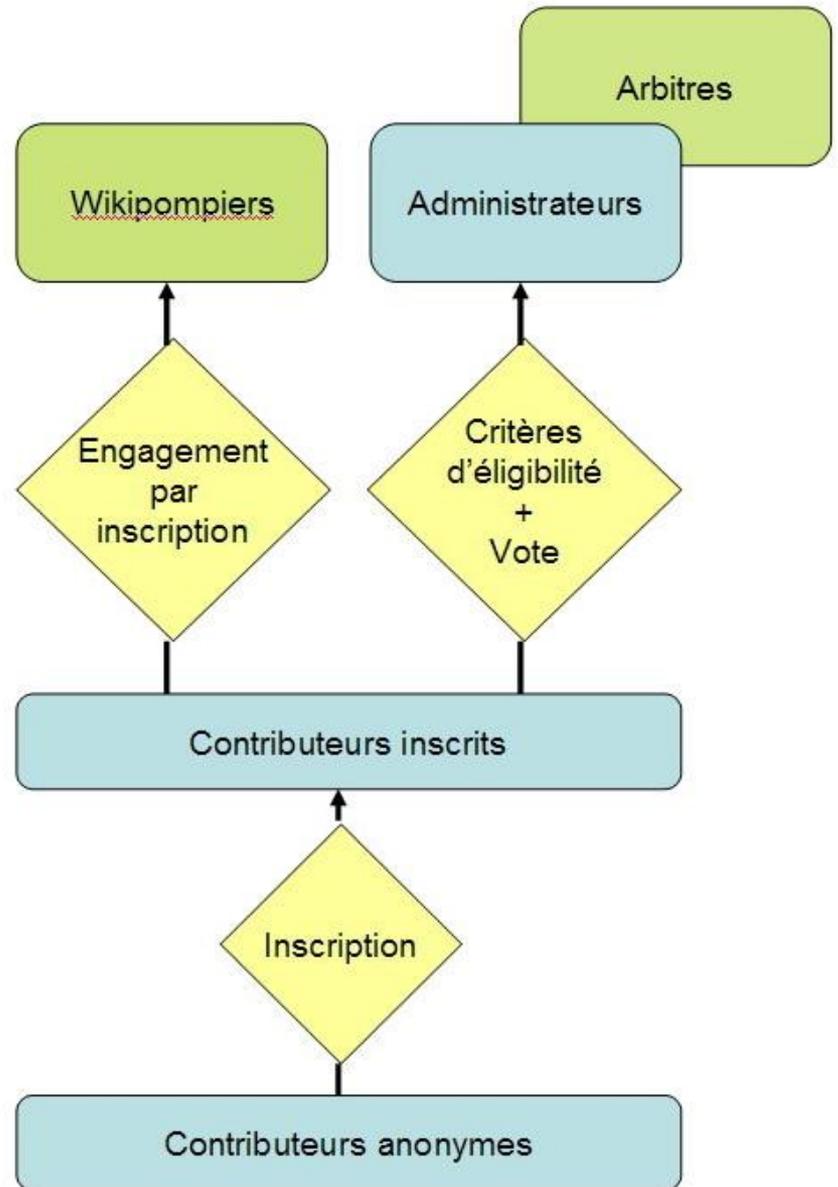
et ses impliqués potentiels (en bleu)

Nota Bene

wikipompiers (éteindre les feux, conflits d'édition, supprimés en 2009)

Salon de médiation

http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Le_salon_de_m%C3%A9diation



Contributeurs anonymes et enregistrés

L'anonyme (l'IP) vs le contributeur enregistré

- Il semble qu'une des IPs anonymes qui faisaient des réversions sauvages se soit maintenant créé un user name fallacieux : "Service de Nettoyage"... quelle prétention, pour juste faire des guerres d'édition sans intervenir en discussion pour justifier ses choix... Sera il nécessaire de remonter la discussion plus haut ? (Psychanalyse)
- Ce serait venu d'une IP j'aurais sans doute effacé à peu près toute son intervention pour "manque de sources" ; comme ça vient d'un vieux de la vieille (ah ce biais envers les contributeurs respectés...) je m'interroge à voix haute. (Bogdanoff)

La communication médiée par les réseaux (CMR)

Premiers travaux dans les années 1990

De nouvelles situations de communication ont fait surgir un écrit interactif et dialogique, qui dispute à l'oral le domaine de l'expression spontanée et de la réaction immédiate. (Anis 1998 :269).

Différents axes de comparaison

- Communication offline vs online (e.g. Baron sur les e-mails par rapport aux courriers)
- Distinctions oral vs écrit vs multimodal
- Communication synchrone vs asynchrone
- Communication privée vs publique
- ...

Les interactions Wikipédia du point de vue des critères de la CMR

Communication **écrite, asynchrone et publique**

- "Cette permanence, liée à l'accessibilité, permet à la fois l'extériorisation et le partage de la cognition, au même titre que lorsque deux personnes rédigent ensemble un texte sur le même ordinateur" (Mangenot, 2002)
- Toute discussion est publique et lisible par tous les participants.
 - Problèmes de face potentiels
- Les pages n'ont pas de version définitive
- La discussion est asynchrone, mais dans une période courte. L'attente d'une réponse peut varier de quelques minutes à plusieurs jours.

Particularité de l'adresse dans les interactions Wikipédia

Adresses fréquentes à la non-personne

- *Floréal a supprimé mes paragraphes discrètement et sans rien justifié, mais aussi le bandeau de non neutralité justifié ici (Féminisme)*

Plusieurs facteurs:

- La discussion autour d'un article étant optionnelle et asynchrone, un contributeur ne peut être certain que son message sera lu
 - Un moyen plus direct d'adresser un message à un contributeur particulier étant de lui écrire directement sur sa page utilisateur mais on change de situation communicative
- Interactions ouvertes et publiques pouvant engager de nombreux contributeurs potentiels (ambiguïté de la seconde personne)

Wikipédia comme corpus

DÉVELOPPEMENT ET EXPLORATIONS DU CORPUS
WIKICONFLITS

Wikipédia comme terrain de recherche: premiers projets

Coopération, conflits, querelles dans le Wikipédia francophone

- AURAY, Nicolas, HURAUULT-PLANTET, Martine, POUDAT, Céline et JACQUEMIN, Bernard (2009),
- JACQUEMIN, Bernard, LAUF, Aurélien, POUDAT, Céline, HURAUULT-PLANTET, Martine et AURAY, Nicolas (2008a et b).
- AURAY, Nicolas, POUDAT, Céline et PONS, Pascal (2007).

Étude de la représentation des sciences dans Wikipédia

- POUDAT, Céline et LOISEAU, Sylvain (2007). « Représentation et caractérisation lexicale des sciences dans Wikipédia » in Tutin, A. (coord.), « Lexique de la langue scientifique », *Revue Française de Linguistique Appliquée*, XII, 2007-2, pp. 29-44.

La question du conflit

Comment définir un conflit?

- **Situation**, dans laquelle des tenants d'un point de vue se confrontent aux supporters de points de vue divergents
- **Processus argumentatif** avec différents rôles concernant une thèse
 - Un conflit suppose au moins **3 mouvements argumentatifs**:
 - (1) assertion d'une proposition;
 - (2) attaques ou doutes exprimés concernant la proposition;
 - (3) défense d'une proposition avec une justification ou une autre attaque.

Marqueurs linguistiques

- Marqueurs de subjectivité et de polarité (personnes, lexique, structures syntaxiques)

Différents types de conflits

Conflits personnels ou affectifs remettant en question la légitimité du contributeur (arguments *ad hominem*)

Conflits méta-argumentatifs ou procéduraux remettant en question le respect des règles du débat, l'organisation du groupe;

Controverses, ayant une dimension sociale, et impliquant des arguments de plus grande échelle correspondant à des discours sociaux

Les types de conflits dans Wikipédia

Les discussions et les conflits dans Wikipédia

- Processus normal: nombreuses modifications avant l'obtention d'articles consensuels
- En anglais: près de 99% des articles seraient élaborés de manière collaborative / pour le reste, des articles polémiques avec des discussions agressives (Auray et al., 2009; Yasseri et al., 2012)

Problèmes de **neutralité**, désaccord de neutralité (NPOV)

- Biais idéologiques, commerciaux, articles trop subjectifs...

Problèmes de **pertinence** (*relevance*), désaccord de pertinence

- Information non pertinente, qualité et style, hors topic, problèmes de références...

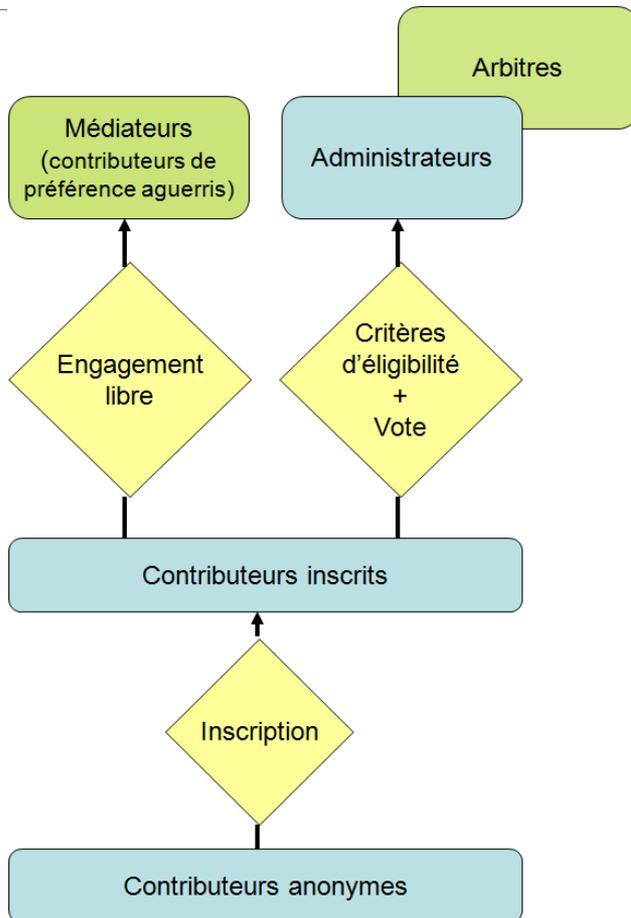
Les conflits dans Wikipédia

Qu'est-ce qu'un conflit dans Wikipédia et comment l'observer?

Du point de vue de Wikipédia

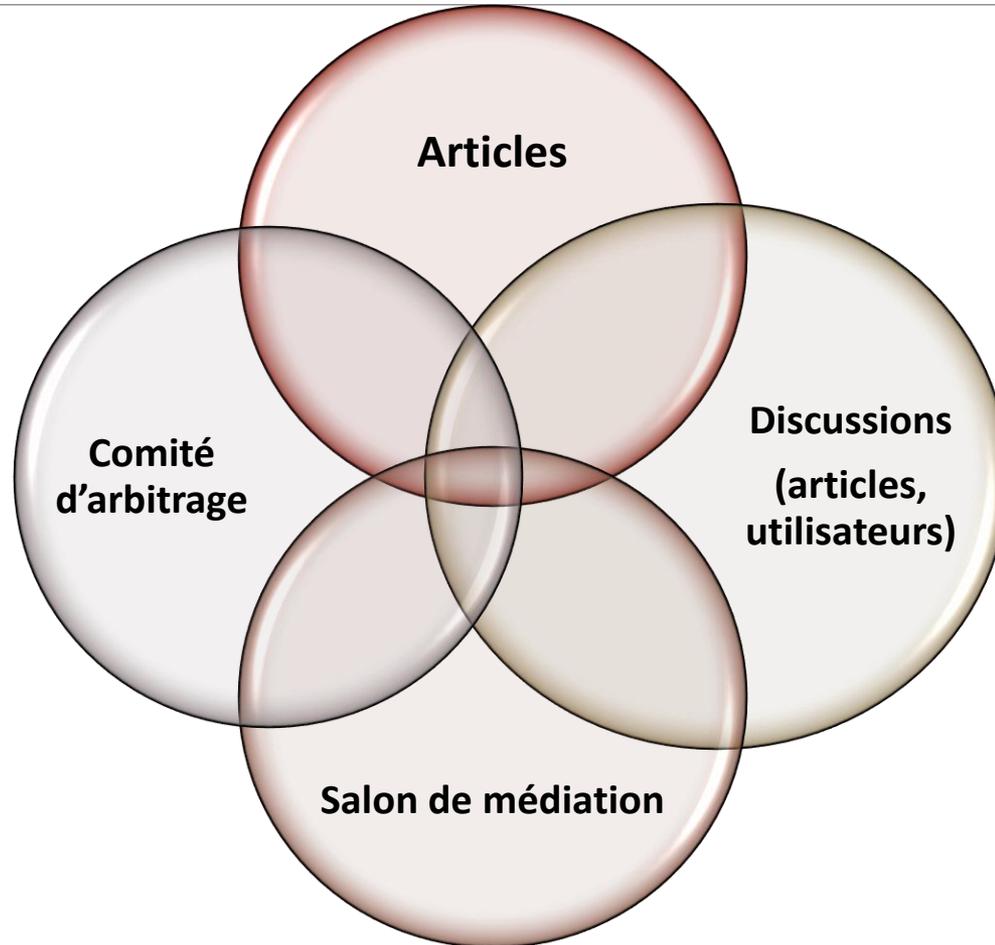
- Identification des conflits par les contributeurs qui les font remonter et appliquent les procédures mises en place
- Un conflit est un blocage susceptible d'entraver l'édition des articles, schématiquement:
 - Conflits et guerres d'éditions (éditions antagonistes et révocations successives) – pages articles;
 - Discussions bloquées ou tendues (NPoV, Pertinence, Appel au calme, médiation) – pages de discussion

Procédures de résolution des conflits dans Wikipédia



- i. Articles
 - Guerres d'édition (révocations / *reverts*)
 - Protections et bandeaux
- ii. Discussions
 - Pages ouvertes par les bandeaux
 - Appel au calme
 - Médiations
- iii. Arbitrages / jugements

Espace d'interaction



Wikiconflits / Projet Coméré

Sous-projet du projet national Coméré (Communication Médiée par les Réseaux)

- Consortium TGIR Huma-num Corpus écrits, 2013-2015
- GT7 Nouvelles formes de communication, coord. Chanier&Poudat
- Principes: **variété, open access, standards**
- Groupe *Nouvelles collectes* (Wikipedia et Twitter, coord. Poudat)

Une perspective linguistique et CMC

Wikiconflits

CONSTITUTION (POUDAT,
GRABAR ET PALOQUE 2017)

Sélection des pages de discussion

- Conflits identifiés par Wikipédia (NPoV, pertinence, protection, médiation...)
- Domaine des sciences et techniques

Extraction

Structuration et annotation

7 ARTICLES + HISTORIQUES +
DISCUSSIONS

Scientificité et pseudo-sciences

- Chiropratique
- Quotient Intellectuel
- Psychanalyse

Controverses sociales / technosciences

- Organisme génétiquement modifié
- Eolienne

Légitimité scientifique et académique d'individus

- Igor et Grichka Bogdanoff

Histoire et épistémologie d'une discipline

- Histoire de la Logique

**4,456 posts rédigés par 3,971 contributeurs, soit
489,000 tokens – 330 Mo au format ZIP**

Pages de discussion à l'étude

article associé	nb. pages**	nb. fils	nb. posts	nb. mots
Chiropratique	2 (+1)	38	285	47 414
Quotient Intellectuel	1 (+0)	45	169	22 815
Psychanalyse	7 (+1)	136	729	80 912
OGM	8 (+1)	373	3 173	345 482
Eolienne	1 (+1)	43	145	15 792
Igor et Grichka Bogdanoff	3 (+1)	129	727	137 809
Histoire de la Logique	1 (+0)	9	45	3 850
Vladimir Poutine*	1 (+1)	68	359	50 664
Attentats du 11 sept. 2001*	3 (+1)	134	1 149	135 103
DISCOWIKI	27 (+7)	975	6 781	839 841

comptage réalisé sur les pages t.q. extraites automatiquement dans WikiTalk

* nouvelles par rapport au corpus WikiConflicts

** page courante et archivées (+ page parallèle c12/11/2021 - SÉMINAIRE CLLE - TOULOUSE ralite)

Wiki Démo

Ouvrage accepté chez Benjamins: *Wikipedia as Corpus. Linguistic corpus building, exploration and analysis* (coord. Poudat, Lungen et Herzberg), à paraître 2022

8 discussions, dans trois langues (français, anglais, allemand)

- Refugee crisis
- Chiropratique
- Vladimir Poutine
- 9/11 attacks
- Psychoanalyse
- OGM
- Feminism
- Gaming discussion

Traitement du corpus

Pré-traitement du corpus

Particularités de la syntaxe wiki (pas de bouton send comme dans les forums): comment segmenter les posts?

- Historique des révisions (Daxenberger, Ferschke and Gurevych 2013)
- **Segmentation de la page discussion et correction manuelle de la syntaxe**

Méthode

- Restitution des dates et des contributeurs (quand c'était possible)
- Correction des références des posts
- Conversion en format XML-TEI (CMC-Core – TEI feature request en cours de finalisation)

pages parallèles >

Autres discussions [liste]
Suppression - Neutralité - Droit d'auteur - Article de qualité - Bon article - Lumière sur - À faire - Archives

méta-données >
- portail
- avancement

Avancement Importance pour le projet

B	Maximum
	Élevée
	Faible

- États-Unis** (discussion • critères • liste • stats • hist. • comité)
- New York** (discussion • critères • liste • stats • hist. • comité)
- Histoire** (discussion • critères • liste • stats • hist. • comité)
- Politique** (discussion • critères • liste • stats • hist. • comité)
- Criminologie** (discussion • critères • liste • stats • hist. • comité)
- Wikipédia 1.0/Les plus consultés** (discussion • critères • liste • stats • hist. • comité)
- Sélection transversale** (discussion • critères • liste • stats • hist. • comité)
- Aéronautique** (discussion • critères • liste • stats • hist. • comité)
- Islam** (discussion • critères • liste • stats • hist. • comité)

liste des fils >

- Sommaire** [masquer]
- Réactions de la défense aérienne
 - le titre
 - Théorie du complot
 - Enquête sur les faux passeports des terroristes en Afghanistan: précisions passeports retrouvé dans les décombres

archivage des anciens fils >

Archives



Page 1, Page 2

titre d'un fil >

> Wikipédia se ridiculise [modifier le code]

post niveau 1 >

> La dernière phrase de l'introduction est tirée d'une vidéo Dailymotion sur un talk show people de Fox News « Certaines personnes ou groupes de personnes contestent la version officielle... » Wikipédia se ridiculise en s'abaissant à ce niveau sur un sujet comme le 11 septembre.

post niveau 2 >

> +1. Pas besoin de se faire l'écho en introduction de théories complotistes fantaisistes et minoritaires. Surtout en renvoyant vers un article aussi PoV que [Théories du complot à propos des attentats du 11 septembre 2001](#) et en sourçant avec une vidéo FoxNews hébergée sur Dailymotion (j'espère par ailleurs qu'avoir simplement donné cet avis ne me vaudra pas un blocage, hein). SM ** ようこそ ** 21 novembre 2010 à 12:18 (CET)

post niveau 3 >

> Qu'on y ajoute du crédit ou pas (et personnellement, je trouve que ces théories sont toutes plus ridicules les unes que les autres), il n'en reste pas moins qu'il existe un gros battage médiatique autour de ces "théories" et qu'un nombre non négligeable de personnes semblent les considérer plausibles. Au point que c'est devenu un indéniable phénomène de société. Ce qu'il fait qu'il est difficile que Wikipedia les passe sous silence. Au demeurant, une section de l'article y est consacrée avec renvoi vers "l'article" ad hoc. C'est sur cette considération que j'avais [modifié](#) la phrase mise en introduction tout en restant réservé par rapport à la référence que je n'avais pas pu consulter à ce moment. L'ayant vue entre-temps, il me semble qu'il faut conserver la phrase en l'état (elle ne fait de rendre compte d'une réalité - les théories du complot - sans prendre position) tout en supprimant la référence (il n'est du reste pas requis de sourcer ce qui figure dans une introduction), qui est loin d'être ce qui ce fait de mieux en la matière. --Lebob (d) 21 novembre 2010 à 13:03 (CET)

post niveau 4 >

> J'entends bien. Tu as parfaitement raison, Wikipédia ne peut pas ignorer l'existence des théories du complot, et la section qui y est consacrée a toute sa place. Simplement, je ne suis pas du tout convaincu par la pertinence d'en parler dès l'introduction. Encore moins avec la source actuelle qu'il faut, dans tous les cas, retirer. SM ** ようこそ ** 21 novembre 2010 à 13:54 (CET)

post niveau 5 >

> Bien d'accord sur la question de la source: d'une part il n'est pas exigé de mettre de sources en introduction et, d'autre part, celle-là est particulièrement mauvaise. Pour le reste, je suggère d'attendre d'autres avis sur le maintien (ou pas) de cette phrase ou d'une possible reformulation d'icelle. --Lebob (d) 21 novembre 2010 à 14:03 (CET)

post niveau 6 >

> ✓ Source retirée, sans toucher au reste pour l'instant, faute de consensus sur ce qu'il convient de faire de cette phrase. SM ** ようこそ ** 21 novembre 2010 à 14:12 (CET)

L'unité d'analyse: le fil de discussion

En dehors de quelques exceptions, le fil de discussion délimite une interaction

- Divisions de la discussion
- Limites des spécificités du format Wiki

Annotation des fils de discussion conflictuels / non conflictuels (Denis et al. 2012 sur la page de discussion de l'article *Pluton*)

- Marqueurs de seconde personne
- Longueur des fils – il faut au moins trois tours de parole pour qu'il y ait conflit, et cinq tours de parole pour qu'on puisse parler de résolution

Message, **fil**, article

Annotation des fils conflictuels

Pré-typage des fils et recherche d'indices conflictuels

Pré-typage des fils:

- **Présence ou non d'un conflit**
- Fil monologal / **dialogal**

Recherche des marqueurs de violence verbale : **actes de langage dépréciatifs directs** (insultes, menaces, mépris, provocations...)

- Insultes="0" (aucun marqueur)
- Insultes="1" (= 1 marqueur)
- Insultes="2" (< 1 marqueur)

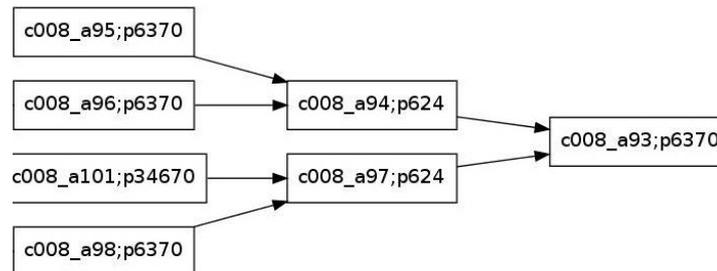
Intervention ou non d'un **médiateur**

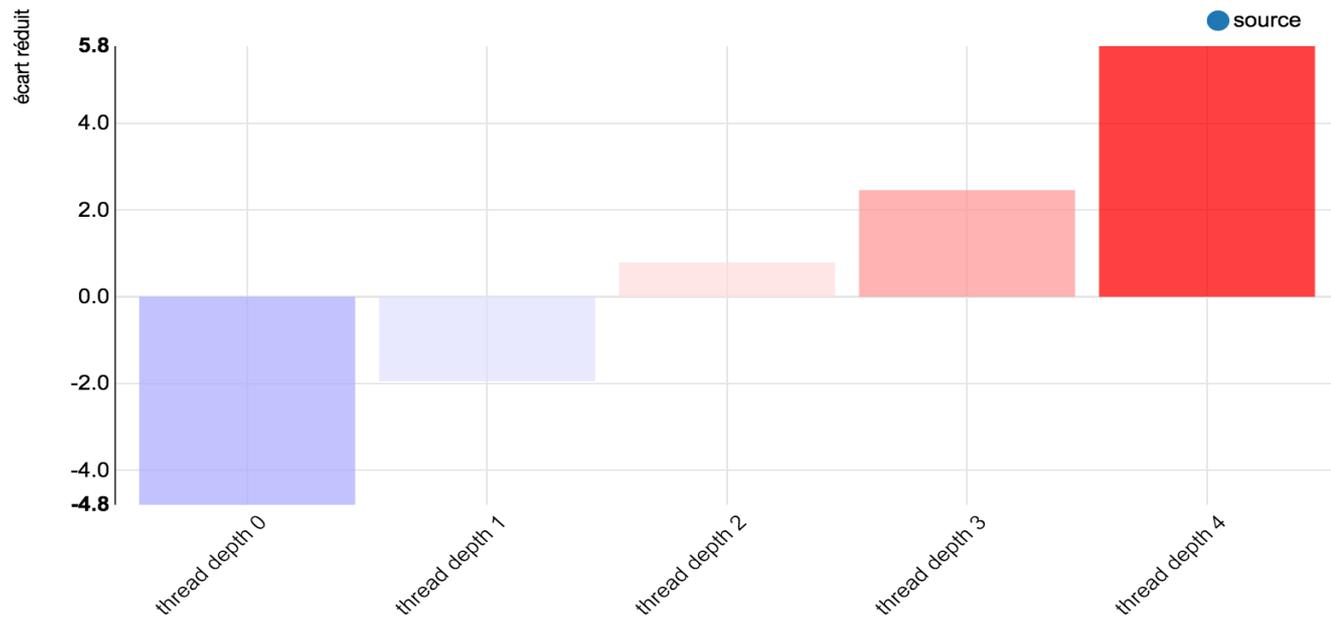
« Bonsoir, J'ai bloqué 3 jours Dujo pour passage en force. Il n'est pas acceptable que le travail collaboratif de discussion qui a lieu ici soit perturbé par des renouveaux de guerre d'édition. Je vous souhaite de continuer à débattre dans le respect mutuel, comme suggéré par ConradMayhew à 7:43. Bien cordialement, »

Comment explorer un corpus d'interactions?

Quelques expériences typologiques

- Fils conflictuels vs non conflictuels (variable binaire, Denis et al. 2012) (Poudat et al. 2016); variable à trois modalités (Poudat 2018)
- Caractéristiques du fil de discussion, e.g.
 - **Longueur**: position du message dans le déroulé du fil / comparaison des fils de longueur identique
 - **Profondeur**: degré de réaction du message / comparaison de fils de profondeur identique
 - Exemple de fil de longueur 3 et de profondeur 4





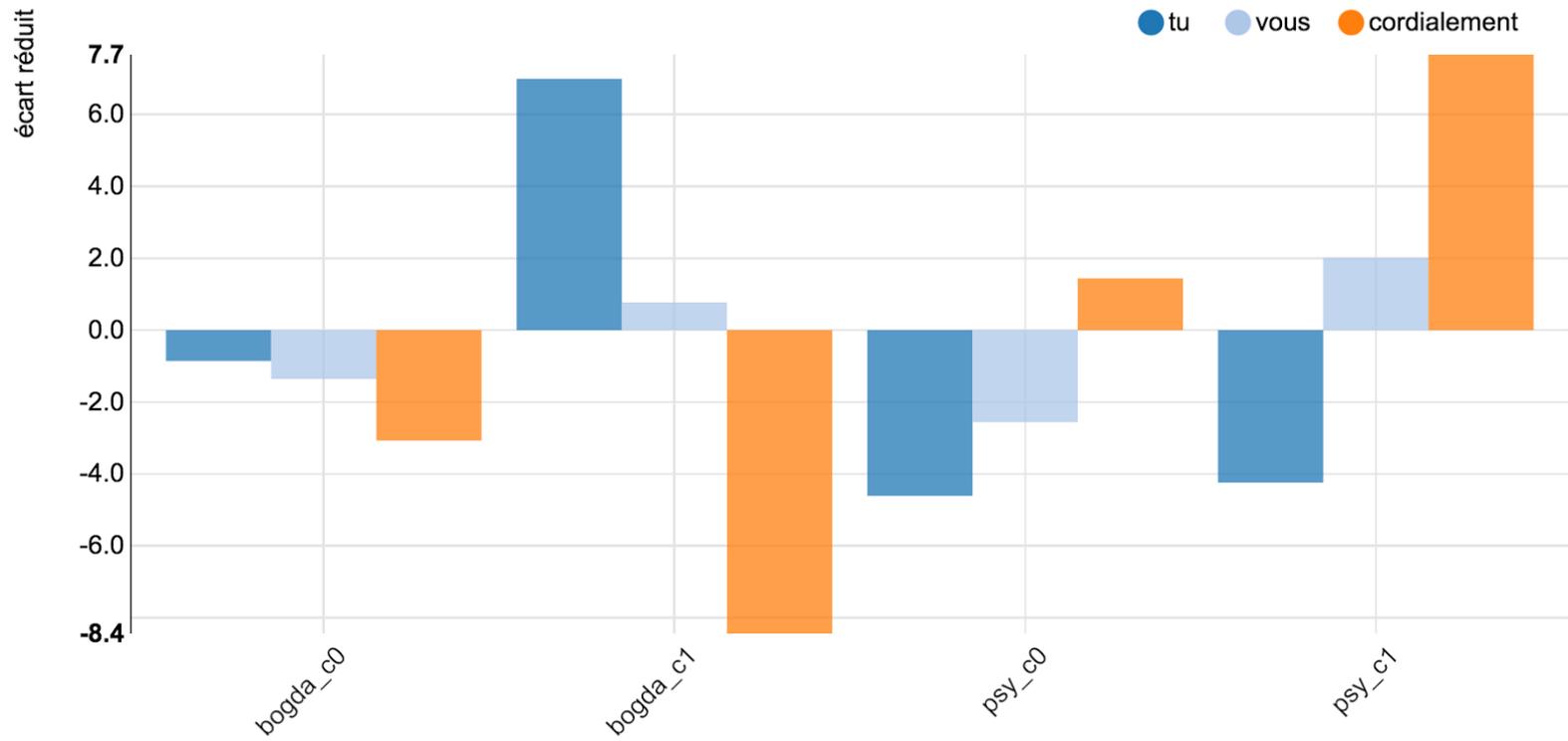
Différents types de marqueurs conflictuels (Poudat et al. 2016)

DISCUSSION BOGDANOFF

écart	corpus	partition	mot
15.98	146	131	Alain
13.12	137	113	frère
12.07	415	243	leur
10.27	101	80	CNRS
10.8	397	224	scientifique
9.75	105	80	université
9.12	97	73	Igor
7.56	46	39	émission
7.9	58	47	doctorat
7.28	118	75	polémique
6.28	61	43	affaire
6.48	34	29	directeur
6.99	184	103	tu
6.52	39	32	notoriété
6	38	30	médiatique
6.13	120	70	thèse
6.89	39	33	Marianne
5.11	133	70	personne
5.13	102	57	site
5.94	42	32	journaliste
5.97	40	31	chercheur
5.04	21	18	jury
5	76	45	physique

DISCUSSION PSYCHANALYSE

écart	corpus	partition	mot
22.42	2490	1774	@card@
20.64	915	742	psychanalyse
12.34	216	192	février
10.54	2217	1381	(
10.46	2248	1397)
9.78	89	86	catégorie
8.13	497	344	--
8.78	148	126	psychanalyste
8.37	480	336	critique
7.98	109	95	noir
7.49	138	113	—
7.66	153	124	cordialement
6.81	53	50	escroquerie
6.62	74	65	décembre
6.52	2449	1425	
5.84	700	438	source
5.55	87	70	psychanalytique
5.46	185	132	vue
5.57	122	93	mars
5.04	201	139	livre
5.03	53	45	psychanalyser
5.24	68	56	freudien
5.21	103	79	fiable



Saisir et décrire les interactions conflictuelles dans Wikipédia

SYSTÈMES DE MARQUEURS LINGUISTIQUES DU
CONFLIT DANS WIKIPÉDIA

Quels marqueurs du conflit?

Quelles sont les particularités linguistiques des discussions conflictuelles ?

Avec quels systèmes de descripteurs en rendre compte ?

- Distinguer désaccords et attaques (Face Threatening Acts)
- Comment les Wikipédiens expriment-t-ils leur désaccord face à un autre contributeur?
- Quels sont les attaques les plus récurrentes dans le corpus?

⇒ Développement de deux systèmes de descripteurs et leurs modèles d'annotation

- Convocation des théories de la (socio-)pragmatique e.g. Goffman 1955, Brown & Levinson 1978, Watts 2003, Roulet et al. 1985, Kerbrat-Orecchioni 1990-2016, Bousfield 2008 / Fracchiolla et al. 2013, Laforest & Vincent 2011, Vincent et al. 2012, Vincent 2010 pour la violence verbale

Annotation systématique du corpus

- 3 annotateurs experts + un groupe d'étudiants de Master
- Recherche de consensus – validations croisées du corpus

1. Expression linguistique du désaccord

Désaccords dans Wikipédia

Un acte de langage réactif qui exprime une réaction négative relative à une assertion, une prise de position préalablement exprimée ou une action (e.g. la modification d'un article).

- Cette réaction négative ne saurait en revanche s'instancier dans une agression, une disqualification, un reproche ou un dénigrement (= attaque)

Un désaccord est une forme d'affrontement verbal qui ne saurait être vue comme négative *a priori*

- la perception et l'évaluation d'un désaccord dépendent sans aucun doute des normes des genres et des cultures.
- Dans Wikipédia, le désaccord est admis et on reconnaît facilement son caractère inévitable (éditions antagonistes d'articles, désaccords sur le fond, la structure des articles)

Désaccords dans Wikipédia

Le désaccord est une **menace potentielle** pour la face des contributeurs (enchaînement non préféré)

- La face est entendue ici comme l'autorité, l'expertise, la légitimité des croyances d'un contributeur sur un sujet donné
- L'interlocuteur ne s'aligne pas sur le locuteur et prend donc le risque de l'opposition et du conflit

Un désaccord qui met du temps à se régler au fil de la discussion a plus de chances de déboucher sur un conflit (répétition d'affrontements verbaux, montée de la tension)

Critères généraux

Formes adressées

Désaccord avec un locuteur

- Inutile de saturer la PdD pour répondre, surtout si tu y mets du hors-sujet par rapport à l'article ! **Je ne suis pas d'accord avec toi** : Nature ou pas, tous les journaux scientifiques ont un comité de lecture, et leurs publications sur le QI sont contradictoires. (QI)

Désaccord avec une proposition ou une action d'un locuteur

- **Je reste par ailleurs en désaccord éditorial** sur la place disproportionnée accordée à l'historique et la non évocation totale d'un état de l'art. (Psychoanalyse)
- **Je ne suis pas d'accord avec** le plan proposé par Flop, avec le tien non plus. (OGM)

Modèle d'annotation

Attributs	Valeurs	Exemples
polarité	accord	je suis d'accord
	désaccord	Je ne suis pas d'accord
type	direct	Accord direct : je suis d'accord, je suis pour X, favorable à X, tout à fait de votre avis, je suis de ton avis, OK pour X... Désaccord direct : je ne suis pas d'accord, en désaccord, je ne suis pas favorable, je suis contre, totalement contre
	indirect	Valeurs illocutoires (diapo suivantes)
atténuation	oui / non	Atténuation d'un accord direct (rare) : je suis assez d'accord Atténuation d'un désaccord direct : Nous sommes en désaccord (mineur) sur un point (mineur)

Expression indirecte du désaccord

Mise en scène d'une sphère d'accord consensuelle par le biais de la concession – accord sur fond de désaccord

Ex. Vos deux premiers arguments sont parfaitement recevables, mais le troisième ne l'est pas : c'est justement parce que cette affaire n'a pas été jugée qu'il y a un problème de démocratie. S'il s'agissait d'un simple problème de corruption jugé, il n'y aurait pas de problème.

Adoucissement du désaccord

Connecteurs d'opposition et de mise à distance: *mais, au contraire, à l'opposé, maintenant*

Restriction (temporaire) de l'annotation aux actes de langage explicites accompagnant un connecteur d'opposition

Expression indirecte du désaccord

Expression d'une opinion contraire

B: **Je ne pense pas que l'on puisse qualifier cela de publicité déguisée**, si le lien est suivi de la mention 'paiement requis' (ou équivalent). Ceci étant, je pense qu'une simple mention du numéro du journal et de sa date de parution sont suffisants pour retrouver l'article, en ligne ou en version papier. -- AlNo 17 janvier 2007 à 12:10 (CET) (Poutine)

Accompagnateurs: marqueurs exprimant la différenciation

- Marqueurs de points de vue: *personnellement, pour moi, de mon point de vue*
- Adverbes énonciatifs de phrase: *franchement, sincèrement*

Expression indirecte du désaccord

Focus sur l'énonciateur et ses défaillances potentielles

Incompréhension

- **Je continue à ne pas comprendre** le rapport entre cet article et les vidéos dont tu postes le lien. (Attentats du 11/09)
- **J'avoue ne pas comprendre** non plus pourquoi ces ajouts documentés ont été supprimés et non transférés dans la page sur la polémique. (Bogdanoff)
- **non, vous m'avez mal lu.** je propose de présenter simplement les faits incontestables (et donc incontestés). (Attentats du 11/09)

Doute

- **je ne suis pas sûr, je ne suis pas certain**

Émotion de type surprise, étonnement

- **un point m'étonne cependant**, en maints endroits, dont ici, on présente Breuer comme faisant partie intégrante de la psychanalyse, alors qu'il n'a jamais été psychanalyste... (Psychanalyse)
- **Autre point qui m'a choqué** sur cette page de discussion, le fait qu'une encyclopédie doit accorder plus de place aux rapports officiels qu'aux interrogations issus de sources non-gouvernementales. (Attentats du 11/09)

Expression indirecte du désaccord

Accord sur fond de désaccord	Concession Forme fréquente 18%	Seuls les accords directs avec une concession (oui mais...) ont été annotés <i>D'accord pour refuser le paragraphe ajouté à partir d'arkiv ; en revanche la suppression de la participation d'AR à la mission ne me semblait pas déraisonnable</i>
Se poser comme source évaluative	Incompréhension Forme fréquente 31%	<i>je ne comprends pas très bien pourquoi il faudrait mentionner leur doctorat dès l'introduction.</i>
	Opinion Forme fréquente 21%	<i>Personnellement / Pour ma part, je ne pense pas, je ne crois pas que...</i> <i>ici je ne pense pas que ce soit judicieux</i>
	Doute Forme rare 4,5%	<i>je doute fort que les médecins étrangers de pays modernes (USA, UK, Allemagne, etc.) soient plus mauvais et moins bien formés que les médecins Français</i>

2. Marqueurs du conflit

ATTAQUES (FTAS)

Conflits et violence verbale

D'après (Auger et Moïse, 2005), la violence verbale est un processus interactionnel qui représente :

- *une montée en tension contextualisée qui se décline à travers différentes étapes (incompréhension, négociation, évitement, renchérissement, renforcement...), marquée par des « déclencheurs de conflit » et par l'emploi d'une importante variété d'actes de langage (harcèlement, mépris, déni, insulte...).*

la violence verbale serait liée aux principes de comportement ou d'action fixés par des individus, des communautés ou des groupes

Trois types de violence verbale (Moïse 2001 ; Auger et al., 2008)

la violence fulgurante correspond à l'augmentation d'une tension dans une conversation qui se manifeste par la présence des indicateurs de conflit, des actes de langage péjoratifs ou des marqueurs discursifs de rupture.

la violence polémique s'appuie sur des actes de langage indirects et implicites ainsi que des figures rhétoriques à but argumentatif.

la violence détournée apparaît dans des échanges faussement aimables et coopératifs où tout compliment possède un sens caché inverse de celui du départ. Ce type de violence a pour objectif de manipuler une personne.

L'insulte

D'après Kerbrat-Orecchioni (1980), l'insulte figure comme acte social portant des conséquences différentes, notamment un conflit. L'insulte elle-même se manifeste par des expressions à connotation négative, autrement dit, des marques axiologiques négatives.

Comment reconnaître un terme insultant et quelles sont ses propriétés ?

- *je vous emmerde espèce de grosse vache stalinienne.*

Acte de langage représentant une adresse *qui témoigne d'un conflit ou le provoque* (Lagorgette, 1994)

L'insulte - Laforest et Vincent (2004)

Toute langue possède ses propres *insultes usuelles* (Perret, 1968) considérées comme dépréciatives par un certain groupe de personnes.

D'une manière générale, ces insultes se basent sur les métaphores et comparent un individu à des animaux ou à des objets perçus négativement.

L'insulte est produite par la première personne (insulteur) et adressée à la seconde personne (insulté). Elle peut réaliser un acte car elle possède un pouvoir performatif.

Il existe deux fonctions d'insulte : rituelle et personnelle.

Les actes de condamnation

Pour (Laforest et Moïse, 2013), l'insulte et le reproche font partie de la catégorie des **actes de condamnation**.

Ils peuvent apparaître dans le même discours et emporter des conséquences similaires

Deux catégories qui m'ont semblé adaptées

- **actes de condamnation du faire** à l'aide desquels l'émetteur exprime son mécontentement vis-à-vis du comportement inapproprié d'un individu présent ou non dans l'interaction.
- **actes de condamnation de l'être** permettent au locuteur d'exprimer son mécontentement à propos d'un trait d'un individu présent ou non dans l'interaction (op.cit).

Le reproche

Acte de parole qui sert à juger négativement le comportement d'un individu et à revendiquer son changement - Kerbrat-Orecchioni (1986)

- une réaction négative du récepteur vis-à-vis du reproche mène souvent au conflit.

Le reproche constitue le plus souvent l'acte directeur d'un énoncé. Il est en général à l'initiative d'un dialogue ou d'un échange, et vient en réaction à un comportement ou une attitude jugés répréhensibles (Denoyelle 2006)

Critères généraux

Conflits adressés

Marqueur de 2nde personne (Denis et al. 2012)

- Dépréciation ou catégorisation dépréciative de l'allocutaire
 - Je pense que vous devez aller vous reposer un minimum. (attentats du 11/09)
 - Je pense que vous continuez à avoir une attitude de blocage contraire aux règles de Wikipédia. (OGM)
 - Je crois que votre point de vue n'est pas le meilleur qui soit, ni le plus exact. (psychanalyse)
 - je crois que votre cas est tout simplement désespéré. (attentats du 11/09)
 - Votre connaissance de Wikipédia est en effet bien basse. (attentats du 11/09)
 - votre comportement nous rappelle celui de certains fanatiques (Bogdanoff)

Modèle d'annotation

Attributes	Values	Examples
type	Être	Et si vous n'étiez si minable, vous vous seriez tu pour ne pas vous ridiculiser davantage. (Féminisme)
	Faire	Encore faut-il que cette décision soit fondée. Tu ne t'es pas justifié sur le fond de la décision. (page Psychanalyse)
	ND	Vous etes malade Pégase? ça doit etre la turista, c'est normal pour quelqu'un qui passe ici en touriste. ça se soigne remarquez, suffit de ne plus y venir. (page Féminisme)
atténuation	oui / non	Malgré mon a priori favorable pour toute personne parlant l'Hindi, je suis au regret de vous dire que vous choisissez mal vos exemples. (page Poutine)

Modèle d'annotation

Attributs	Valeurs	Exemples
ironie	oui / non	Je sais bien que tu veux sauver le monde de ces infâmes crapules que sont les jumeaux, mais il serait temps de percuter que tu brasses de l'air depuis des mois et des mois (page Bogdanoff)
Valeur illocutoire	reproche	Une fois de plus tu essaies de changer le sujet pour éviter de répondre à la question posée. (page Chiropratique)
	menace	Et pire, sans référence sérieuse, ce que tu as enlevé sera rétabli, car ces personnages sont des universitaires, donc leurs publications sont des sources entrant dans les critères de wp. (page QI)
	insulte	si tu n'as rien de nouveau à présenter, on va arrêter de perdre notre temps avec un <u>vrai croyant</u> [lien vers la page du Syndrome du vrai croyant] qui continue à pleurnicher et à faire semblant de pas voir les réponses des autres (page Chiropratique)

Premiers éléments de résultats

600 attaques annotées dans le corpus

La plupart des attaques sont des reproches

Les insultes et les catégorisations dépréciatives sont peu fréquentes

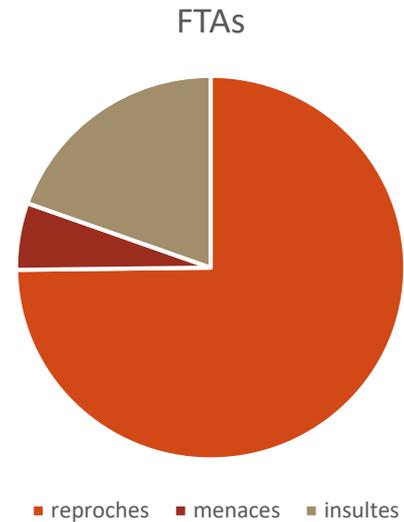
Les condamnations du faire sont beaucoup plus fréquentes que les condamnations de l'être

- 85% des reproches sont des condamnations du faire
- 42% des insultes condamnent l'être du contributeur

Des analyses sont en cours pour déterminer si l'ironie (présente dans un peu moins de 40% des attaques) adoucit ou intensifie l'attaque (humour vexatoire)

La frontière entre reproches et insultes est parfois ténue

Nécessité de revoir les relations entre les deux systèmes (désaccord indirect / reproche)



Campagne de crowdsourcing

Évaluation et amélioration de l'annotation – stage de Roxane Bois

Un premier bêta test mené sur une petite cohorte de 20 personnes montre que les trois actes de langage reproches / insultes / menaces posent des problèmes d'identification

- Manque d'éducation à l'identification des actes de langage négatifs?

Cadre de la campagne

- Étape de certification
- 9 annotations par séquence annotée

Perspectives

Approfondissement et stabilisation de la typologie

Application potentielle: content moderation, projet Wikimedia?

Analyse de la séquentialité des fils

Étude en cours des méthodes d'annotation et d'exploration des corpus annotés (nouveau projet CORLI)

Extension à l'anglais dans une perspective contrastive

Extension à d'autres genres de la CMC, à commencer par les forums

Références

Bousfield, D. (2008). *Impoliteness in Interaction*. Amsterdam: John Benjamins Publishing Company.

Brandes, U. and Lerner, J., 2007: Revision and co-revision in wikipedia: Detecting clusters of interest. In *Proceedings of International Workshop Bridging the Gap Between Semantic Web and Web 2.0*, 4th European Semantic Web Conference (ESWC'07), Innsbruck, Austria.

Brown, P., & Levinson, S. (1978, 1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press. [First published 1978 as part of Ester N. Goody (Ed.): *Questions and Politeness*.]

Denis, A., Quignard, M., Fréard, D., Détienne, F., Baker, M., Barcellini, F., 2012: Détection de conflits dans les communautés épistémiques en ligne. In *Actes de la Conférence sur le Traitement Automatique des Langues Naturelles-2012*.

Ferschke, O., Gurevych, I., Chebotar, Y., 2012: Behind the article: Recognizing dialog acts in wikipedia talk pages. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pages 777-786. Association for Computational Linguistics.

Fracchiolla, B., Romain, C., Moïse, C., Auger, N., & Collectif. (2013). *Violences verbales?: Analyses, enjeux et perspectives*. Rennes: PU Rennes.

Goffman, E. (1955). On Face-Work: An analysis of ritual elements in social interaction, *Psychiatry: Journal of Interpersonal Relations*, 18(3), pp. 213–231.

Goffman, E. (1961). *Encounter: Two studies in the sociology of interaction*. Oxford, England: Bobbs-Merrill.

Ho-Dac, L.-M. and Laippala, V., 2017: Le corpus WikiDisc, une ressource pour la caractérisation des discussions en ligne. In Wigham, C.R., Ledegen, G. (eds) *Corpus de communication médiée par les réseaux : construction, structuration, analyse*. Collection Humanités Numériques. Paris : L'Harmattan. pp. 107-124.

Kerbrat-Orecchioni, C. (1990-1992-1994). *Les interactions verbales*, t. I-II-III, Paris, Armand Colin.

Kerbrat-Orecchioni, C. (1995). *Variations culturelles et universaux dans les systèmes conversationnels: Aspects cognitifs et sociaux de l'interaction* (pp. 97–111). Presented at the analyse des interactions. Colloque international. Retrieved from <http://cat.inist.fr/?aModele=afficheN&cpsidt=3295668>.

Kerbrat-Orecchioni, C. (2005). *Le discours en interaction*, Paris Armand Colin.

Kerbrat-Orecchioni, C. (2016). *Les actes de langage dans le discours - Théorie et fonctionnement*. Paris: Armand Colin.

Références

- Kittur, A. and Kraut, R. E., 2008: Harnessing the wisdom of crowds in wikipedia: quality through coordination. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, pages 37-46. ACM.
- Kittur, A. and Kraut, R. E., 2010: Beyond wikipedia: coordination and conflict in online production groups. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, pages 215-224. ACM.
- Laforest, M., & Vincent, D. (2011). La qualification péjorative dans tous ses états. *Langue française*, 144, pp. 59–81.
- Moïse, C., Meunier, E., & Romain, C. (2015). *La violence verbale dans l'espace de travail?: Analyses et solutions*. Levallois-Perret: BREAL.
- Poudat, C., & Landragin, F. (2017). Explorer un corpus textuel : Méthodes - pratiques - outils. De Boeck supérieur.
- Poudat, C. and Loiseau, S., 2007: Représentation et caractérisation lexicale des sciences dans wikipédia. *Revue française de linguistique appliquée*, 12(2):29-44.
- POUDAT, Céline, GRABAR, Natalia, PALOQUE-BERGES, Camille, CHANIER, Thierry et KUN, Jin (2017). « Wikiconflits : un corpus de discussions éditoriales conflictuelles du Wikipédia francophone » in Wigham, C.R & Ledegen, G., *Corpus de communication médiée par les réseaux : construction, structuration, analyse*. Collection Humanités numériques. Paris : L'Harmattan.
- Poudat, C., Vanni, L., Grabar, N., 2016: How to explore conflicts in french wikipedia talk pages? In JADT, pages 645-656.
- Roulet, E. et al. (1985). *L'articulation du discours en français contemporain*. Berne: P. Lang.
- Stvilia, B., Twidale, M. B., Smith, L. C., Gasser, L., 2008: Information quality work organization in wikipedia. *Journal of the American Society for Information Science and Technology*, 59(6):983-1001, April.
- Viégas, F. B., Wattenberg, M., Dave, K., 2004: Studying cooperation and conflict between authors with history flow visualizations. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 575-582. ACM.
- Viegas, F., Wattenberg, M., Kriss, J., van Ham, F., 2007: Talk Before You Type: Coordination in Wikipedia. In *40th Annual Hawaii International Conference on System Sciences, 2007. HICSS 2007*, pages 78-78, January.
- Vincent, D., & Bernard Barbeau, G. (2012). Insulte, disqualification, persuasion et tropes communicationnels: à qui l'insulte profite-t-elle? *Argumentation et Analyse du Discours*, 8. Retrieved from <http://aad.revues.org/1252>.



Saisir l'expression linguistique du conflit dans les discussions Wikipédia. Corpus et systèmes de marqueurs

CÉLINE POUDAT

CELINE.POUDAT@UNIV-COTEDAZUR.FR