

20^e conférence

TALN

15^e édition

RECITAL

17 au 21 Juin

2013

LES SABLES D'OLONNE

www.taln2013.org



Vers une détection automatique des chiasmes

Détecter les chiasmes dont les termes principaux entretiennent une forte proximité morphologique ou sémantique

MARIE DUBREMETZ

université
Paris Ouest

Nanterre La Défense



& UPPSALA
UNIVERSITET

Les Sables d'Olonne, jeudi 20 juin 2013

Plan de la présentation

- 1 Introduction : les différents types de chiasmes
- 2 Problématique et hypothèses
- 3 État de l'art
- 4 Démarche
 - Test de l'hypothèse sur les contraintes
 - Test de l'hypothèse sur les outils de TAL et les chiasmes
- 5 Bilan

Une définition vague

Définition traditionnelle

Figure formée de deux groupes d'éléments linguistiques dont l'ordre est inverse. Exemple : Bonnet blanc et blanc bonnet.



Figure: Schéma définitoire du chiasme

Une définition vague

Définition traditionnelle

Figure formée de deux groupes d'éléments linguistiques dont l'ordre est inverse. Exemple : Bonnet blanc et blanc bonnet.

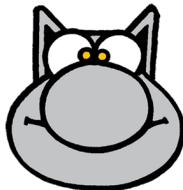
Bonnet blanc
X
blanc bonnet

Figure: Schéma définitoire du chiasme

Introduction » types de chiasmes

Question linguistique

Qu'est-ce qu'un chiasme de mots ?



La réponse en image

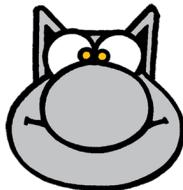
L'exemple de la Bande-Dessinée



Introduction » types de chiasmes

Question linguistique

Qu'est-ce qu'un chiasme de mots ?



La réponse en image

L'exemple de la Bande-Dessinée



Exemple d'antimétabole

L'antimétabole

Figure de style jouant sur la réversion de mots strictement identiques [Dupriez, 2003]



Exemple de chiasmes jouant sur la morphologie

Le chiasme flexionnel

Figure de style jouant sur la réversion de mots qui diffèrent par leur flexion



Exemple de chiasmes jouant sur la morphologie

Le chiasme dérivationnel

Figure de style jouant sur la réversion de mots de même racine

"Les Réunionnais sont des Français à part entière... sinon ils sont entièrement à part."

Exemple de chiasmes jouant sur des mots sémantiquement proches

L'antimétalepse [Diderot et D'Alembert, 1789]

Figure de style jouant sur la réversion de mots sémantiquement proches



Question linguistique

Qu'est-ce qu'un chiasme de mots ?

La réponse du linguiste

Le chiasme est une figure de style...

- 1 Regroupant plusieurs phénomènes linguistiques
- 2 Efficace car agit sur la forme (lexico-syntaxique) voire sur le fond !
- 3 à potentiel : persuasif, humoristique...

Question linguistique

Qu'est-ce qu'un chiasme de mots ?

La réponse du linguiste

Le chiasme est une figure de style...

- ➊ Regroupant plusieurs phénomènes linguistiques
- ➋ Efficace car agit sur la forme (lexico-syntaxique) voire sur le fond !
- ➌ à potentiel : persuasif, humoristique...

Introduction » Problématique (transition)

Question linguistique

Qu'est-ce qu'un chiasme de mots ?



La réponse du chat

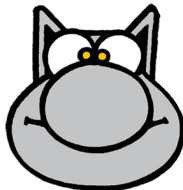
C'est un trait d'esprit...



Introduction » Problématique (transition)

Question linguistique

Qu'est-ce qu'un chiasme de mots ?



La réponse du chat

C'est un trait d'esprit...



Problématique : comment identifier tous ces types de chiasmes automatiquement ?

- Hypothèse 1 : L'application de contraintes simples et formelles est nécessaire
- Hypothèse 2 : Les outils de TAL peuvent nous aider à couvrir tous les chiasmes de mots

Les recherches sur l'antimétabole

- [Hromada, 2011] : Reconnaître trois paires de mots en réversion (Regex)
 - Très précis
 - Beaucoup d'oublis
- [Gawryjolek, 2009] : extraire toutes les doubles paires de mots incluses sans exception
 - 100% de rappel
 - moins de 2% de précision

Démarche empirique

À quoi ressemble typiquement un chiasme ?

À deux propositions cohérentes sur le fond comme sur la forme :

*Le Parti socialiste est un **parti sans leader**.*

*François Bayrou est un **leader sans parti**.*

Qu'est-ce qui caractérise généralement un faux positif ?

- L'incohérence dans la ponctuation

futur géant " (Le Monde) Journal :

*Le **Monde** (França) Titre : ' Comment Lula gère un **futur***

- L'incohérence dans les termes principaux utilisés

*Il n'empêche qu'avant **ou après nous nous** rencontrions à Paris , à Berlin **ou** ailleurs*

Conclusion de l'hypothèse sur les contraintes

Résultat

Stopwords + ponctuations vérifiées = précision $\times 32$

Conclusion

Par l'ajout de contraintes très formelles on peut augmenter significativement la précision de détection des antimétaboles

Test de l'hypothèse sur les outils de TAL et les chiasmes

1. test sur un texte préalablement lemmatisé automatiquement

- outils : programme Python + TTG + Flemm
- résultats : satisfaisants

2. Test sur un texte préalablement normalisé à la racine

- outils : idem que 1. + Snowball
- résultats : mauvais

3. Test sur texte préalablement lemmatisé avec une fonction de comparaison sémantique

- outils : idem que 1. + thésaurus Open Office
- résultats : mitigés mais prometteurs



Test de l'hypothèse sur les outils de TAL et les chiasmes

Conclusion : les outils de TAL nous ont-ils aidé à trouver plus de chiasmes ?

Les ressources TAL en Français que nous avons testées permettent de retrouver antimétaboles, chiasmes flexionnels et chiasmes portant sur les liens de synonymie uniquement.

Que nous a apporté cette recherche ?

- Une définition synthétique, adaptée pour le TAL
- Un texte annoté en dépit du manque d'exemples des ouvrages de référence
- Un test des solutions existantes
- Travail d'un sujet peu traité, dialogue interdisciplinaire...

Ouverture

- Des progrès à faire en précision
- Envisager le parsing ?
- Traiter les autres types de chiasmes

Que nous a apporté cette recherche ?

- Une définition synthétique, adaptée pour le TAL
- Un texte annoté en dépit du manque d'exemples des ouvrages de référence
- Un test des solutions existantes
- Travail d'un sujet peu traité, dialogue interdisciplinaire...

Ouverture

- Des progrès à faire en précision
- Envisager le parsing ?
- Traiter les autres types de chiasmes

-  DIDEROT, D. et D'ALEMBERT, J. I. R. (1789).
Antimétabole, Antithèse, Antimétalepse.
-  DUPRIEZ, B. (2003).
Gradus, les procédés littéraires.
Union Générale d'Éditions 10/18.
-  GAWRYJOLEK, J. J. (2009).
Automated Annotation and Visualization of Rhetorical Figures.

Master thesis, University of Waterloo.
-  HROMADA, D. D. (2011).
Initial Experiments with Multilingual Extraction of Rhetoric
Figures by means of PERL-compatible Regular Expressions.
In RANLP, Hissar, Bulgaria.

Une dernière ?



Des questions ?

- <http://stp.lingfil.uu.se/~marie/chiasme.htm>
- marie.dubremetz@lingfil.uu.se



Une dernière ?



Des questions ?

- <http://stp.lingfil.uu.se/~marie/chiasme.htm>
- marie.dubremetz@lingfil.uu.se

