Université de Paris

Wmatrix, Sketch Engine et Laughter : méthode semi-automatique pour détecter la prosodie sémantique dans les *TED Talks*

Anastasia Buturlakina
sous la direction de Natalie Kübler, Professeur

Problématique

- Prosodie sémantique : difficile à détecter – intuition linguistique, lecture des concordances
- Transgression humoristique de la PS : marqueurs du rire

→ Proposer une méthode semi-automatique de détection de la prosodie sémantique dans le corpus des *TED Talks* anglais

Plan

- Problématique
- Approche théorique :
  - prosodie sémantique
  - *TED Talks*
  - humour dans les *TED Talks*
- Corpus
- Méthodologie
- Analyses
- Discussions
- Perspectives

Prosodie sémantique : définitions

- Prosodie sémantique : tendance statistiquement significative d'une unité lexicale ou d'une expression à se trouver dans un environnement comportant une évaluation, d'entrer en co-occurrence avec les mots ayant une connotation
- Connotation : sens évaluatif associé à l'unité lexicale qui lui est inhérent et dont les locuteurs d'une langue sont conscients

	Sens inhérent	**Phénomène collocationnel**
Prosodie sémantique	–	+
Connotation	+	+ / –

Prosodie sémantique : principes

- « cachée », mais accessible dans le corpus grâce aux concordances
- « graduée », continuum autour d'une prosodie sémantique neutre avec des polarités négative et positive graduées (Partington, 2004)
- Transgression de la PS habituelle ou co-occurrence de plusieurs mots/expressions à PS opposées → effet humoristique
- Positif / négatif : ce qui a un impact positif / négatif pour les humains (Hunston, 2007 : 253, 263)

TED Talks

- TED : *Technology, Entertainment, Design,*
- Accessibles
- Double audience
- Communication entre les spécialistes et les non-spécialistes
- Vulgarisation
- Genre hybride : oral, mais préparé à l'écrit et révisé
- Traits caractéristiques : la narration, l'évaluation et l'humour (Scotto di Carlo, 2015)

Humour dans les *TED Talks*

L'humour dans les *TED Talks* sert à créer la réaction émotionnelle chez le public, à établir la connexion avec lui, à divertir, à attirer l'attention et à la retenir (Scotto di Carlo, 2014).

- Les marqueurs du rire du public dans les retranscriptions : « (Laughter) »

Corpus des *TED Talks*

- Documents : 2 533 retranscriptions des *TED Talks* anglais
- Dates : De juin 2007 à mars 2018
- Taille : 4 429 042 de tokens
- URL : https://www.ted.com/

Méthodologie

Transgression humoristique de la prosodie sémantique : unité lexicale / expression à PS (X) + mot(s) à connotation inattendue (ou dans l'ordre inverse) → marqueur du rire

- Annotation sémantique USAS de Wmatrix pour détecter les mots à connotation : les émotions connotées (« + », « - »)
- Formule pour détecter la transgression humoristique de la prosodie sémantique :

  X + « mot d'émotion » + marqueur du rire dans une portée de 5 mots

- Requête CQL dans Sketch Engine :

  [lempos=".*E.*"] []{0,5} [word="\("] [word="[Ll]augh.*"]

➢ 1900 concordances extraites → analyse qualitative pour détecter l'incongruité linguistique

Annotation et corpus

- Wmatrix : textes
  - étiquetage morphosyntaxique (CLAWS)
  - étiquetage sémantique (USAS)
  - non adapté à la consultation des corpus

  Fonction → annotation du corpus
- Sketch Engine : corpus
  - consultation du corpus pré-annoté par Wmatrix dans Sketch Engine
    - requête CQL pour détecter la transgression humoristique de la prosodie sémantique
  - Consultation des corpus généraux EnTenTen et BNC pour valider la prosodie sémantique des lexies /expressions
    - Word Sketch, collocations et fréquences → validation des candidats PS

Résultats

- Transgression humoristique de la prosodie sémantique détectée
- Autres stratégies linguistiques pour faire de l'humour dans les *TED Talks*
- Plusieurs stratégies linguistiques combinées renforcent l'effet humoristique produit

→ Humour linguistique : transgression du principe de l'idiome de Sinclair

Stratégies linguistiques derrière l'humour dans les *TED Talks*

a) transgression de la prosodie sémantique d'un mot ou d'une expression ;
b) transgression de la prosodie sémantique d'un mot en tant que fonction pragmatique ;
c) conflit entre les prosodies sémantiques opposées de plusieurs mots ou expressions ;
d) conflit entre les connotations opposées de plusieurs mots ou expressions ;
e) transgression du principe idiomatique ;
f) oxymore ;
g) transgression de la préférence sémantique ;
h) transgression du parallélisme ;
i) conflit des registres ;
j) conflit des domaines ;
k) remplacement d'un élément de l'expression figée par l'antonyme ;
l) retournement d'un concept à PS positive en le décrivant avec un concept à PS négative.

Exemples

1. faced with such outrageous hyperbole
2. grim inspiration
3. reputable criminal with good ratings
4. lose 37 years of emotional baggage
5. experience the visceral feeling of trusting strangers
6. created artificial stupidity
7. enjoy it [obituary]
8. revel in putridity
9. precious little cannibal
10. former baby
11. screams good taste
12. oh beloved dentist
13. the sediment… map of the seabed radioactivity… favorite sushi
14. may she [Blockbuster Video] rest in peace
15. enforced rest
16. waging peace
17. make peace with those dogs
18. horrible creative-destructive annoying
19. intriguing… great, thundering twit
20. cheery about failing
21. suffering from childhood
22. unfortunately, successful
23. dirty work of happiness
24. global warming fun
25. big happy, nervous, uncomfortable family
26. get the loss
27. a guardian angel… the Panic Monster

	Exemples																										
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
Stratégies linguistiques de l'humour																											
a																											
b																											
c																											
d																											
e																											
f																											
g																											
h																											
i																											
j																											
k																											
l																											

Analyses

Now let me tell you, if you want to experience the visceral (E1) feeling (E1) of trusting (E6+) strangers – (Laughter) I recommend this, especially if those strangers are drunk German people.

Connotation positive : trust

Connotations négatives : visceral, stranger

} conflit entre les connotations

PS positive : trust

PS négatives : visceral, experience

} conflit entre les prosodies sémantiques

Préférence sémantique : trust + X digne de confiance

négation + trust + anyone / anybody

→« Enchâssement » évaluatif (Partington, 2017) :

experience [the [visceral feeling] of (trusting) [strangers]]

Analyses

Or this ad that I read in the New York Times: "Wearing a fine watch speaks loudly of your rank in society. Buying it from us screams good taste (... O4.2 E2+ ...)." (Laughter)

PS négative : scream

PS positive : good taste

} conflit entre les prosodies sémantiques

PS neutre : speak (loudly) of

Connotation négative : scream

Connotation positive : good taste

} conflit entre les connotations

Connotation neutre : speak (loudly) of

→ Parallélisme entre les deux phrases : une phrase au gérondif sert de sujet, suivi par un verbe et un objet

→ Hypothèse : parallélisme transgressé

Discussions

+ Méthode de détection de la transgression humoristique de la PS est efficace
+ Plusieurs autres stratégies linguistiques créent l'effet humoristique dans les *TED Talks*
+ Ces stratégies linguistiques interviennent à plusieurs, renforçant l'effet humoristique produit
+ Humour linguistique : transgression du principe de l'idiome de Sinclair
– Méthodologie adaptée seulement aux retranscriptions des genres oraux avec les marqueurs du rire
– Lecture des concordances chronophage

➢ Nouvelles prosodies sémantiques

➢ Lexies / expressions à PS négatives détectées sont légèrement plus nombreuses que celles à PS positives

Perspectives

- Étude traductologique et contrastive de la transgression humoristique de la prosodie sémantique entre l'anglais et le français :
  - EN original – FR traduit
  - FR original – EN traduit
  - EN original – EN traduit
  - FR original – FR traduit

- Voie pour la détection automatique de la PS utilisant la méthode de l'apprentissage profond
- Étude en profondeur des stratégies linguistiques de l'humour repérées
- Étude en profondeur du rôle de l'humour dans les *TED Talks*

Bibliographie

- Hunston, S. (2007). Semantic prosody revisited. *International Journal of Corpus Linguistics*, 12(2), 249–268.
- Nerhardt, G. (1977). Incongruity and Humour. *Reports from the Department of Psychology, the University of Stockholm.* Stockholm: Department of Psychology, University of Stockholm.
- Partington, A. (2017). Evaluative clash, evaluative cohesion and how we actually read evaluation in texts. *Journal of Pragmatics,* 117, 190-203.
- Partington, A. (2004). "Utterly content in each other's company": semantic prosody and semantic preference. *International Journal of Corpus Linguistics* 9(1), 131–156.
- Rayson, P. (2008). From key words to key semantic domains. *International Journal of Corpus Linguistics*. 13(4), 519-549
- Rayson, P., Archer, D., Piao, S. L., McEnery, T. (2004). The UCREL semantic analysis system. *Proceedings of the workshop on Beyond Named Entity Recognition Semantic labelling for NLP tasks in association with 4th International Conference on Language Resources and Evaluation (LREC 2004),* 25th May 2004, Lisbon, Portugal. Paris: European Language Resources Association, 7-12.
- Scotto di Carlo, G. (2015). Stance in TED talks: Strategic use of subjective adjectives in online popularisation. Ibérica 29 (1), 201–222.
- Scotto di Carlo, G. (2014). Humour in popularization: Analysis of humour-related laughter in TED talks. *The European Journal of Humour Research,* 1(4), 81-93.