

Utiliser les interjections pour les détecter les émotions

Amel Fraise Patrick Paroubek

TALN, Caen, Juin 2015

- 1 Introduction
- 2 Travaux antérieurs
- 3 Création du corpus émotionnel
- 4 Création du lexique d'émotions
- 5 Évaluation
- 6 Conclusion

Microbloggage

- 284 millions d'utilisateurs actifs mensuels
- 500 millions de tweets par jour en moyenne
- Taille de message réduite à 140 caractères
 - Les émoticônes
 - Les interjections

Ça fait VLAM ! ça fait SPLATCH ! et ça
fait CHTUCK !
Ou bien BOMP ! ou HUMPF ! parfois même PFFF !

« Comic Strip », *Serge Gainsbourg*.

- Les interjections onomatopéiques : **interjections primaires émotives** [Kleiber, 2006]
 - Celles qui imitent un bruit naturel qui peut être d'origine humaine (*ouille!*, *ha!*, *ouf!*, *hihi!*, etc.) ou non-humaine (*baoum!*, *tic tac!*, *flic floc!*, etc.) [Barberis, 1992]
- Les interjections non onomatopéiques [Gonçalves, 2008]
 - Ensemble de mots ou expressions figées empruntant leur formulation à d'autres classes de mots :
 - noms (pardon, flûte, sans blague, etc.)
 - verbes (allez!, vive!, etc.)
 - adjectifs (hardi!)
 - adverbes (là, comment, eh bien)

- Elles sont nécessairement les indices linguistiques d'une émotion
- Elles sont donc des marques de subjectivité
- Aucun lien syntaxique ne la relie à son contexte discursif

- Étude des interjections dans le discours : [*Garcia-Fernandez et al., 2010*]
- Utilisation dans les conversations en ligne "Chat" [*Falaise, 2005*]
- Marques modales qui permettent aux locuteurs de faire porter une émotion ou une attitude subjective sur l'énonciation d'un contenu [*Halté, 2013*]
- En fouille d'opinions :
 - Les émoticônes [*Read, 2005 ; Pak & Paroubek, 2010*]
 - Les mots-dièses (Hashtags) [*Mohammad, 2012 ; Qadir & Riloff, 2013 ; Fraise & Paroubek, 2014*]
 - **MAIS jamais les interjections**

- Définition de classes émotionnelles et les interjections liées

#	Classe	Inter. onoma.	Inter. non onoma.
1	COLÈRE	Argh , pff	Voyons !
2	PLAISIR	hihi , haha	lol , youpi
3	PEUR	aah	
4	APAISEMENT	ouf	
5	TRISTESSE	Aïe, ouille	zut, hélas
6	INSATISFACTION	Bof	
7	MÉPRIS	Beurk	
8	SURPRISE NÉGATIVE	oups	

TABLE: Les classes émotionnelles exprimées ayant des interjections clairement définies sur Twitter.

Recherche basée sur les interjections

Interjection	Nombre de messages	% de messages
Argh	1078	5,65
pff	2362	12,39
Voyons!	456	2,39
hihi	3350	17,57
haha	3987	19,14
lol	1456	7,63
youpi	1211	6,35
aah	844	4,42
ouf	1300	6,82
Aïe	320	1,67
ouille	337	1,76
zut	158	0,82
hélas	987	5,17
Bof	650	3,41
beurk	435	2,82
oups	467	2,45
Total tweets	19061	100
Total utilisateurs	19032	

- Information mutuelle : [*Fano, 1961 ; Church & Hanks, 1990*]

$$\text{IM}(X, Y) = \log_2 \left(\frac{P(X, Y)}{P(X) \cdot P(Y)} \right) \quad (1)$$

$$\text{IM}(m, e) = \log_2 \left(\frac{\text{freq}(m, e)}{\text{freq}(m) \cdot \text{freq}(e)} \right) \quad (2)$$

$$\text{IM}(m, \neg e) = \log_2 \left(\frac{\text{freq}(m, \neg e)}{\text{freq}(m) \cdot \text{freq}(\neg e)} \right) \quad (3)$$

$$\text{Asso}(m, e) = \text{IM}(m, e) - \text{IM}(m, \neg e) \quad (4)$$

Avec :

$$\text{freq}(m, e) = \frac{|T_{m,e}|}{|T|} \quad (5)$$

$$\text{freq}(m) = \frac{|T_m|}{|T|} \quad (6)$$

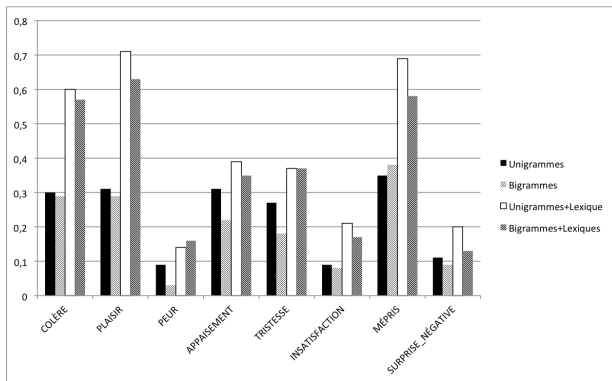
$$\text{freq}(e) = \frac{|T_e|}{|T|} \quad (7)$$

#	Classe émotionnelle	# termes	Exemples
1	COLÈRE	234	colère, furieux, fâché
2	PLAISIR	423	super, enjoy, cool
3	PEUR	184	peur, tremble, inquiet
4	APAISEMENT	178	ouf, enfin, fini
5	TRISTESSE	168	triste, seul, mal
6	INSATISFACTION	58	bof, insatisfait, moyen
7	MÉPRIS	198	déteste, dégoûtant, dégueulasse
8	SURPRISE NÉGATIVE	87	non, stupéfait, impossible

TABLE: Nombre de termes par classe d'émotion dans le lexique.

- Tâche détection d'émotion
- Corpus DOXA [*Paroubek et al., 2010*]
- Apprentissage automatique
 - Machine à vecteur support : SVM
 - Traits : Unigrammes, bigrammes, lexique
 - Poids : TF-IDF normalisé

$$w(t_i) = \frac{tfidf(t_i)}{avg.tf(t_i)} \quad (8)$$



- La performance de classification de classes fréquentes comme *PLAISIR*, *COLÈRE*, *MÉPRIS*, *TRISTESSE* et *DÉGOÛT* est meilleure que celles qui sont moins fréquentes dans les données d'apprentissage comme *PEUR*, *SURPRISE NÉGATIVE* et *INSATISFACTION*. Cependant dans les deux cas, les vecteurs de traits ayant obtenus les meilleurs scores en mesure F1 sont les *unigr.TF-IDF/avg.tf+lexique* et *bigr.TF-IDF/avg.tf+lexique*.

Classes	unigrammes			unigr.+lexique			Δ	bigrammes			bigr.+lexique			Δ
	P.	R.	F1	P.	R.	F1		P.	R.	F1	P.	R.	F1	
COLÈRE	0,32	0,29	0,3	0,49	0,48	0,48	+0,18	0,29	0,31	0,29	0,56	0,46	0,5	+0,21
PLAISIR	0,39	0,27	0,28	0,56	0,51	0,53	+0,25	0,28	0,31	0,29	0,51	0,41	0,45	+0,19
PEUR	0,11	0,09	0,09	0,17	0,12	0,14	+0,05	0,07	0,02	0,03	0,13	0,21	0,16	+0,13
APAISEMENT	0,35	0,28	0,31	0,41	0,39	0,39	+0,08	0,29	0,18	0,2	0,35	0,37	0,35	+0,15
TRISTESSE	0,29	0,27	0,27	0,39	0,36	0,37	+0,1	0,17	0,21	0,18	0,41	0,34	0,37	+0,19
INSATISFACTION	0,11	0,09	0,09	0,17	0,29	0,21	+0,12	0,1	0,07	0,08	0,13	0,27	0,17	+0,09
MÉPRIS	0,41	0,32	0,35	0,47	0,42	0,44	+0,09	0,39	0,38	0,38	0,61	0,56	0,58	+0,2
SURPRISE NÉGAT.	0,13	0,11	0,11	0,23	0,19	0,20	+0,09	0,12	0,08	0,09	0,18	0,11	0,13	+0,04

- Les vecteurs de traits ayant obtenus les meilleurs scores en mesure F1 sont les *unigr.TF-IDF/avg.tf+lexique* et *bigr.TF-IDF/avg.tf+lexique*.
(+0,25%, +0,21%)

- Évaluer et étudier l'apport des interjections pour les systèmes de fouille d'opinions
- Méthode automatique pour collecter et annoter automatiquement des tweets
- Création d'un corpus émotionnel : 19061 tweets repartis sur 8 classes d'émotion
- Lexique affectif : 1530 mots répartis sur les 8 classes d'émotions
- Tâche de détection d'émotion : amélioration de performance des systèmes : de +0,04 à +0,21

