



HAL
open science

L'expression des émotions dans les textes pour enfants : constitution d'un corpus annoté

Aline Étienne, Delphine Battistelli, Gwéno   Lecorv  

► To cite this version:

Aline   tienne, Delphine Battistelli, Gw  no   Lecorv  . L'expression des   motions dans les textes pour enfants: constitution d'un corpus annot  . 6e conf  rence conjointe Journ  es d'  tudes sur la Parole (JEP, 33e   dition), Traitement Automatique des Langues Naturelles (TALN, 27e   dition), Rencontre des   tudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (R  CITAL, 22e   dition), 2020, Nancy, France. pp.205-212. hal-02784769v2

HAL Id: hal-02784769

<https://hal.science/hal-02784769v2>

Submitted on 18 Jun 2020 (v2), last revised 23 Jun 2020 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destin  e au d  p  t et    la diffusion de documents scientifiques de niveau recherche, publi  s ou non,   manant des   tablissements d'enseignement et de recherche fran  ais ou   trangers, des laboratoires publics ou priv  s.

L'expression des émotions dans les textes pour enfants : constitution d'un corpus annoté

Aline Étienne^{1,2} Delphine Battistelli¹ Gwénoél Lecorvé²

(1) Univ. Paris-Nanterre, CNRS, MoDyCo, 200 av. de la République, 92001 Nanterre, France

(2) Univ Rennes, CNRS, IRISA, 6, rue de Kerampont, 22300 Lannion, France

{aline.etienne, delphine.battistelli}@parisnanterre.fr,
gwenole.lecorve@irisa.fr

RÉSUMÉ

Cet article présente une typologie de divers modes d'expression linguistique des émotions, le schéma d'annotation sous Glozz qui implémente cette typologie et un corpus de textes journalistiques pour enfants annoté à l'aide de ce schéma. Ces travaux préliminaires s'insèrent dans le contexte d'une étude relative au développement des capacités langagières des enfants, en particulier de leur capacité à comprendre un texte selon des critères émotionnels.

ABSTRACT

Expressing emotions in texts for children: constitution of an annotated corpus.

This paper presents a classification of various types of linguistic expression of emotions, the Glozz annotation schemata implementing this classification and a corpus of journalistic texts for children manually annotated with this schemata. This preliminary work fits in the broader context of a psycholinguistic study of children's language abilities development, and more precisely of their ability to understand a text depending on emotional criteria.

MOTS-CLÉS : émotions, schéma d'annotation, textes pour enfants.

KEYWORDS: emotions, annotation schemata, texts for children.

1 Introduction

Nous nous intéressons ici aux modes d'expression des émotions dans les textes pour enfants. Ce travail s'inscrit dans l'objectif à plus long terme du développement d'outils à même de repérer automatiquement des portions de textes difficiles à comprendre pour des enfants jeunes lecteurs. Dans la lignée de travaux en psycholinguistique comme ceux de (Zwaan et Radvansky, 1998), nous abordons la compréhension d'un texte comme la construction d'une représentation mentale de la situation décrite par le texte, intégrant plusieurs dimensions situationnelles (le temps, l'espace, la cause par exemple). À l'instar de travaux comme (Blanc, 2010), nous intégrons les émotions ressenties par les personnages (à distinguer des émotions ressenties par le lecteur (Dijkstra et al., 1995)) aux dimensions situationnelles d'un texte, et nous nous intéressons en particulier ici à leurs divers modes d'expression linguistique. Ces émotions sont en effet primordiales pour la compréhension de texte puisqu'elles permettent de saisir pourquoi les personnages agissent de telle ou telle manière et ainsi de (re)créer des liens de causalité entre des situations évoquées dans un texte. De fait, il a été montré que, rendus saillants par des émotions, certains événements (ceux présentant un enjeu pour les personnages), sont mémorisés puis restitués plus aisément, chez les

enfants (Blanc, 2010 ; Davidson et al., 2001), comme chez les adultes (Dijkstra et al., 1995). Le présent article décrit le corpus annoté en émotions constitué afin de mieux cerner l'expression linguistique des émotions dans des textes journalistiques pour enfants. Après un état de l'art du traitement des émotions en psycholinguistique, en linguistique et en TAL (section 2), nous détaillons les étapes de l'annotation manuelle du corpus (section 3). Les premiers résultats d'exploration de ce corpus annoté sont abordés en section 4.

2 Les émotions en psycholinguistique, en linguistique et en TAL

En psycholinguistique, l'impact favorable des émotions sur la compréhension des textes chez les enfants a été démontré, voir notamment (Blanc, 2010) et (Davidson et al., 2001). En tant qu'individus en construction (développement linguistique, cognitif etc.), les enfants ne saisissent cependant pas toujours complètement l'émotion exprimée, ce qui pourrait amoindrir l'effet positif escompté de la présence d'émotions sur l'accès au texte. Il a ainsi été montré qu'avant 10 ans, les émotions de base (colère, dégoût, joie, peur, surprise et tristesse, cf. (Ekman, 1992)) sont mieux retenues que les émotions complexes (culpabilité, embarras, fierté, jalousie, cf. (Lewis, 2008)) (Davidson, 2006 ; Blanc et Quenette, 2017). La façon dont les émotions sont exprimées influence également leur compréhension (Blanc, 2010 ; Creissen et Blanc, 2017). De 6 à 10 ans, les émotions explicitement mentionnées dans un texte, que ce soit *via* des descriptions de comportements associés à une émotion - par exemple "il éclate en sanglots" - ou *via* l'emploi d'unités lexicales désignant des émotions (heureux, effrayer, ...), sont mieux comprises que les émotions simplement suggérées par la situation exposée dans un texte (ex. "Le loup arrive"). Dans l'étude de la compréhension de textes par les enfants, l'expression linguistique des émotions occupe ainsi une place déterminante. Cette question représente par ailleurs un enjeu de taille en linguistique. Le rôle crucial que les émotions jouent dans la communication verbale (ex. persuader son auditoire lors d'une argumentation rhétorique, (Micheli, 2010)) les érige en objet d'étude de choix mais leur omniprésence dans le discours rend leur modélisation linguistique particulièrement complexe (Micheli, 2014). De nombreux travaux se focalisent alors sur l'étude du lexique émotionnel (ex. EmoBase (Diwersky et al., 2014)), délaissant les marqueurs ne relevant pas du lexique des émotions (ex. structures syntaxiques ou discursives). Lorsque ces marqueurs sont pris en compte, les travaux n'abordent le plus souvent que des catégories sémantiques restreintes (ex. analyse des expressions de la peur (Bresson et Dobrovol'skij, 1995) ; analyse des structures (morpho-)syntaxiques exprimant la joie et la rage (Gross, 1995)) ou des structures précises (ex. structures des verbes de sentiments (Mathieu, 2006)). La typologie des "modes de sémiotisation" des émotions de (Micheli, 2014) fait alors figure d'exception, puisque ce travail propose une caractérisation générale (absence de focus sur une catégorie émotionnelle spécifique) des différentes manières d'exprimer les émotions dans les textes destinés aux adultes. Cette typologie distingue les émotions dénotées directement par le lexique (appelées émotions "désignées"), celles que l'on devine grâce à la structure syntaxique d'un énoncé et aux choix des mots opérés par l'énonciateur (appelées émotions "montrées") et celles suggérées par la situation décrite par le texte (appelées émotions "étayées").

Les émotions constituent une catégorie sémantique dont les marqueurs linguistiques, impliquant plusieurs niveaux de la langue (lexique, syntaxe etc.), sont difficiles à circonscrire. Dès lors, le repérage automatique des catégories émotionnelles dans un texte paraît être une tâche délicate. Les quelques travaux de TAL qui s'y attellent ont recours à des lexiques émotionnels et n'intègrent pas, à notre connaissance, de marqueurs linguistiques autres que strictement lexicaux. (Mohammad, 2011) emploie par exemple un lexique émotionnel (le NRC Emotion Lexicon) afin de repérer et de quantifier automatiquement les termes de 8 catégories émotionnelles (colère, dégoût, joie, peur, surprise, tristesse, confiance et anticipation/attente) dans des contes de fées et romans anglophones. Il est à noter par ailleurs que, même si fortement connexes, les travaux sur la détection d'émotions

sont à différencier selon nous des travaux sur le repérage de sentiments (catégorisés principalement en positifs, négatifs ou neutres) ou d'opinions (catégorisés principalement en favorables, défavorables ou neutres). (Hamon et al., 2015) illustrent cette distinction en présentant différents travaux de l'édition 2015 du défi fouille de texte (DEFT), visant notamment à repérer automatiquement dans des tweets francophones des sentiments (satisfaction, insatisfaction), des opinions (accord, valorisation, désaccord, dévalorisation) et des émotions (plaisir, apaisement, amour, surprise positive, déplaisir, dérangement, mépris, surprise négative, peur, colère, ennui, tristesse). Ces travaux ont utilisé des lexiques et des méthodes d'apprentissage automatique. Si nous nous situons dans le champ de la détection d'émotions, des correspondances pourront malgré tout être faites entre les catégories émotionnelles que nous annotons et les polarités négatives et positives, établissant ainsi un pont avec le domaine de l'analyse de sentiments. Enfin, l'annotation des émotions se révèle également être une tâche complexe, comme le montrent (Bostan et Klingler, 2018) en comparant différentes ressources annotées en émotions. Ils mettent ainsi en évidence des divergences aussi bien dans la nature des textes annotés (contes, blogs, actualité ...) que dans le choix des catégories émotionnelles (en lien avec diverses approches psychologique des émotions, notamment celle de (Ekman, 1992)) ou les procédures d'annotations employées (annotation par experts, *crowdsourcing* ...) - diversité à laquelle vient s'ajouter le présent travail.

A ce jour, dans le champ du TAL, aucun travail de caractérisation linguistique fine de l'expression des émotions dans les textes pour enfants ni aucune ressource permettant d'effectuer un tel travail ne semblent avoir été produits. Notre travail vise à combler, en partie du moins, ces lacunes.

3 Constitution du corpus annoté en émotions

Notre méthodologie générale peut être résumée aux trois étapes suivantes : (1) Trouver des pistes théoriques issues de la littérature linguistique et psycholinguistique pour explorer l'expression des émotions dans les textes pour enfants ; (2) Procéder à des analyses linguistiques fines d'extraits de notre corpus afin de vérifier la pertinence des pistes théoriques (catégories émotionnelles, marqueurs linguistiques) ; (3) Systématiser l'étude du corpus *via* un schéma d'annotation. L'objectif de (3) est double : confirmer à l'échelle du corpus entier la pertinence des catégories et des critères retenus d'après (2) ; et obtenir des données quantitatives pour dégager des marqueurs linguistiques diversifiés (i.e. autres que strictement lexicaux) des émotions dans les textes pour enfants. Nous avons étudié un corpus de 97 numéros (octobre 2015 à mars 2019) du journal d'actualité en ligne *Le P'tit Libé*, adressé aux 7-12 ans, et qui vise à expliquer "l'actu des grands" aux enfants. Nous disposons donc de 97 fichiers (un numéro par fichier), soit environ 216K tokens.

3.1 Pistes théoriques explorées et choix opérés

La constitution de notre corpus annoté en émotions vise à rendre compte de la diversité des marqueurs linguistiques des émotions (ex. structures syntaxiques) mais aussi à mettre en évidence des liens entre différentes émotions (ex. apparition conjointe fréquente de la peur et de la tristesse) ou l'utilisation préférentielle de certains modes d'expression pour verbaliser certaines émotions. Comme pour toute tâche d'annotation, nous avons établi un schéma d'annotation en définissant les unités linguistiques à repérer au sein du corpus ainsi que les catégories à leur associer. Notre schéma d'annotation devait permettre le repérage de termes du lexique émotionnel (ex. le mot "triste"), mais aussi de structures syntaxiques (ex. dislocations), de marques typographiques (ex. points d'exclamation) voire de structures textuelles (ex. groupe d'énoncés averbaux). La table 1 reprend et met en regard des travaux de linguistique (Micheli, 2014) et en psycholinguistique

(Blanc, 2010, 2017) sur les modes d'expression linguistique des émotions. En nous appuyant sur ces deux types de travaux, nous avons décidé de retenir la distinction terminologique suivante pour les catégories d'émotions : “désignée”, “montrée”, “étayée” et “comportementale”.

	Catégories			
Micheli (2014)	Émotion « dite » <i>ex. Paul est heureux.</i>		Émotion « montrée » <i>ex. « Ah ! Quel endroit merveilleux », dit Paul.</i>	Émotion « étayée » <i>ex. Après un long voyage, Paul arrive enfin dans la maison de vacances de ses rêves.</i>
Blanc et al. (2010 ; 2017)	Émotion « désignée » <i>ex. Paul est heureux.</i>	« Expression comportementale de l'émotion » <i>ex. Paul sourit.</i>		Émotion « suggérée » <i>ex. Après un long voyage, Paul arrive enfin dans la maison de vacances de ses rêves.</i>
Catégories retenues	Émotion « désignée » <i>ex. Paul est heureux.</i>	Émotion « comportementale » <i>ex. Paul sourit.</i>	Émotion « montrée » <i>ex. « Ah ! Quel endroit merveilleux », dit Paul.</i>	Émotion « étayée » <i>ex. Après un long voyage, Paul arrive enfin dans la maison de vacances de ses rêves.</i>

TABLE 1 : Comparatif des typologies des modes d'expression des émotions et catégories retenues

L'étude linguistique de l'expression des émotions en français souligne l'importance du schéma actanciel (i.e. qui fait quoi à qui) des segments textuels émotionnels (Mathieu, 2006 ; Bresson et al., 1995 ; Micheli, 2014). D'après ces travaux et l'analyse d'extraits de notre corpus, nous avons émis l'hypothèse que l'entité affectée par l'émotion (i.e. l'individu ressentant l'émotion) occupe une place centrale dans l'expression et la compréhension des émotions. Nous avons donc décidé d'annoter ces entités, ainsi que le lien qui les unit aux unités émotionnelles. Selon nous, des informations sur ces entités seront pertinentes pour l'analyse linguistique et doivent figurer dans notre schéma d'annotation : la nature du segment textuel désignant l'individu affecté par l'émotion ; et des caractéristiques de l'individu lui-même, avec les traits [\pm animé,humain] et [\pm collectif] (i.e. emploi générique d'un référent et non utilisation du pluriel). Ces caractéristiques pourraient fournir des pistes pour régler (en partie) le problème de la polysémie comme dans “Marie est isolée”, où “isolée” prend un sens émotionnel (tristesse) grâce au trait [+animé,humain] (vs. “La maison est isolée” où “maison” est [-animé]). Un encodage plus complet du schéma actanciel des émotions nécessiterait une unité pour la cause des émotions (i.e. ce qui déclenche l'émotion chez l'individu). Cependant, d'après l'analyse d'extraits de notre corpus, nous avons supposé que la cause d'une émotion sera moins aisément délimitable que l'entité qui la ressent. Cet élément ne figure donc pas dans notre schéma. Enfin, nous avons inclus une distinction entre les émotions de base vs. complexes, pertinente pour l'étude de la compréhension de textes par les enfants. Afin d'analyser plus finement la réalisation du champ sémantique des émotions, nous avons distingué 10 catégories émotionnelles : 6 émotions “de base” (colère, dégoût, joie, peur, surprise et tristesse) et 4 émotions complexes (culpabilité, embarras, fierté et jalousie), reprises de (Davison, 2006 ; Blanc et Quenette, 2017).

3.2 Schéma d'annotation

Notre schéma aborde les émotions en se focalisant sur le mode d'expression linguistique, avec un type d'unité pour chacun des quatre modes d'expression des émotions et un cinquième pour l'entité qui ressent l'émotion. Des traits associés à chaque type d'unité permettent d'affiner les annotations. Pour les unités émotionnelles, il s'agit du “nom de l'émotion”, soit la catégorie émotionnelle

exprimée par le segment textuel (ex. *peur* pour “apeuré”, “Quelle horreur !” et “mort”) et de la “nature du segment” textuel. Cette expression, volontairement peu précise, vise à englober la variété des marqueurs linguistiques à repérer. Cette information est particulièrement utile pour les émotions “montrées” car cette catégorie très hétérogène inclut aussi bien des mots isolés (ex. interjections), que des structures syntaxiques (ex. dislocations) ou des groupes de phrases (ex. enchaînement d’énoncés averbaux). L’unité “entité” a trois traits : $[\pm \text{ animé, humain}]$; $[\pm \text{ collectif}]$; et “Nature du segment” textuel désignant l’entité ressentant l’émotion. Enfin, notre schéma possède une relation (nommée “Affecte”), typée par le trait “Syntaxe”, servant à relier une unité émotionnelle à l’unité textuelle désignant l’individu affecté par l’émotion.

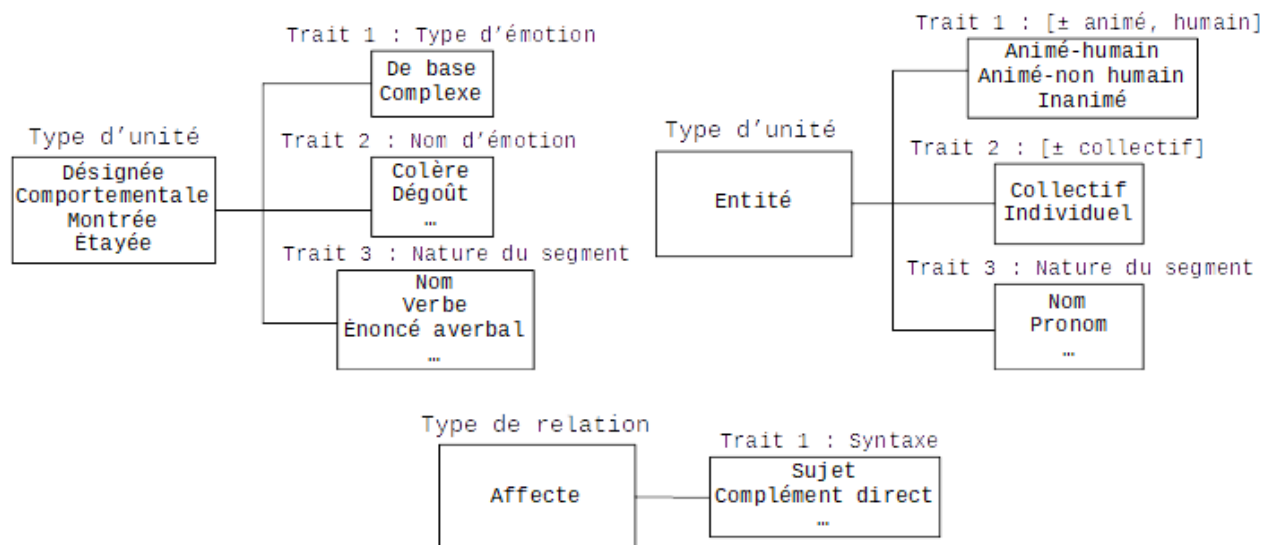


FIGURE 1 : Architecture de notre schéma d’annotation

L’objectif de l’annotation du corpus est de croiser les informations émotionnelles (de base ou complexe, nom de l’émotion) avec d’autres types d’informations (mode d’expression, nature du segment textuel) pour montrer que l’expression des émotions ne repose pas uniquement sur le lexique émotionnel. L’annotation manuelle du corpus a été effectuée *via* la plateforme Glozz (Widlöcher et Mathet, 2012) qui présente plusieurs avantages pour notre travail : (i) grande liberté de schéma d’annotation (traits pour caractériser les unités) ; (ii) diversité des annotations possibles (unités allant du caractère au paragraphe ; chevauchement d’unités ; relations entre unités) ; (iii) pour la suite de nos travaux, calcul d’accord inter-annotateurs aisé grâce à la structure des fichiers Glozz. Une pré-annotation automatique du lexique émotionnel a été opérée par l’application d’une adaptation du lexique EMOTAIX (Piolat & Bannour, 2009) sur le corpus non-lemmatisé *via* un script Python. L’adaptation du lexique a consisté principalement en un remaniement des catégories (conservation des catégories liées à nos 10 émotions ; fusion de catégories proches, ex. *terreur*, *inquiétude*, *angoisse* et *peur* fusionnées en *peur*) et une augmentation en formes fléchies grâce au Lefff (Sagot, 2010) à partir des lemmes du lexique. Cette pré-annotation a ensuite été vérifiée et affinée manuellement, sur Glozz.

4 Exploration du corpus annoté : quelques résultats

Au total, nous avons délimité 2368 unités émotionnelles dans le corpus, dont environ 41,5% d’émotions étayées, 37,1% d’émotions désignées, 11,8% d’émotions comportementales et 9,6% d’émotions montrées. La majorité des unités expriment une émotion de base (89,1%), notamment à cause du nombre plus élevé de catégories d’émotions de base que complexes (6 contre 4). Cependant, une catégorie d’émotion de base est réalisée par 351 unités en moyenne contre 65 pour

une catégorie d'émotion complexe, ces dernières sont donc ici bien moins représentées. Sur l'ensemble des données, l'émotion la plus fréquente est la peur (29,1% des unités). Viennent ensuite la colère (20,9%), la joie (15,2%) et la surprise (13,7%). Les autres catégories représentent moins de 10% des unités (tristesse 9,4%, fierté 5,3%, embarras 4,1%, culpabilité 1,4%, dégoût 0,8% et jalousie 0,1%). La répartition change lorsqu'on regarde les proportions de chaque catégorie émotionnelle selon le mode d'expression (cf. figure 2).

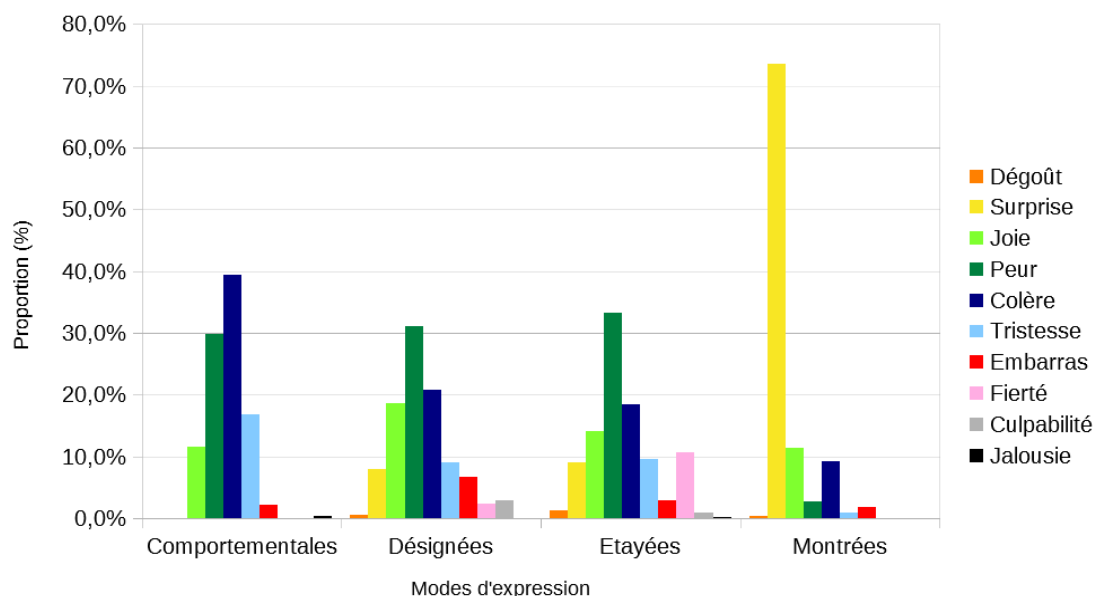


FIGURE 2 : Proportions des catégories émotionnelles selon le mode d'expression

La différence de répartition observée à l'échelle du corpus et en fonction des modes d'expression laisse supposer que des modes d'expression sont utilisés de manière préférentielle pour verbaliser certaines émotions (*ex. montrer la surprise*). Nous avons délimité 1110 unités désignant une entité affectée par une émotion. Il y a moins d'entités que d'unités émotionnelles car soit l'entité n'est pas exprimée, soit plusieurs unités émotionnelles affectent la même entité. Les entités sont réalisées principalement par des pronoms (48,5%) ou des syntagmes nominaux (noms : 37%, noms propres : 11,3%) puis quelques déterminants (3,2%) et un adjectif (dans *à la surprise générale*). Les émotions du corpus affectent surtout des humains (97,3%) et des entités collectives (59,1%).

L'annotation manuelle du corpus a souligné diverses difficultés inhérentes à notre tâche : repérer l'entité qui ressent l'émotion ; identifier précisément l'émotion en jeu parmi les 10 retenues. L'annotation ayant été réalisée par une seule personne, il sera en outre nécessaire de faire une campagne d'annotation pour évaluer la pertinence du schéma d'annotation et la faisabilité de la tâche par un calcul d'accord inter-annotateurs, selon un protocole encore à définir. Nous souhaiterions par ailleurs étendre notre corpus à d'autres genres textuels (encyclopédie, roman).

5 Conclusion

Nous avons constitué une ressource jusqu'ici absente : un corpus de 97 textes journalistiques pour enfants annotés en 10 types d'émotions et selon divers modes d'expression linguistiques mobilisés, soit au total 2368 unités émotionnelles de taille et de nature diverse annotées. Ce corpus annoté contribuera à des travaux dans trois domaines : en linguistique, pour l'exploration de la diversité de l'expression des émotions, en dehors du lexique stricto sensu ; en TAL, pour des systèmes d'apprentissage automatique ; en psycholinguistique, comme matériau de test au sein d'un

protocole expérimental visant à évaluer la compréhension de différents types d'émotions auprès de populations d'enfants, tâche que nous poursuivons actuellement au sein du projet ANR TextToKids. De plus, nous envisageons plusieurs pistes pour approfondir notre étude préliminaire, notamment une comparaison avec l'expression des émotions dans des textes non destinés aux enfants, la mise en regard de l'expression des émotions selon les différentes thématiques des articles et selon leurs auteurs ou encore l'enrichissement de notre schéma d'annotation en intégrant la cause des émotions.

Remerciements

Ce travail a bénéficié du soutien financier de l'Agence Nationale de la Recherche (ANR) dans le cadre du projet TREMoLo (ANR-16-CE23-0019) et du projet TextToKids (ANR AAPG 2019).

Références

- BLANC, N. (2010). La compréhension des contes entre 5 et 7 ans: Quelle représentation des informations émotionnelles?. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 64(4), 256–265. DOI : [10.1037/a0021283](https://doi.org/10.1037/a0021283).
- BLANC, N., & QUENETTE, G. (2017). La production d'inférences émotionnelles entre 8 et 10 ans : quelle méthodologie pour quels résultats ? *Enfance*, 4(4), 503-511. DOI : [10.4074/S0013754517004141](https://doi.org/10.4074/S0013754517004141).
- BOSTAN L.-A.-M. & KLINGER R. (2018, August). An analysis of annotated corpora for emotion classification in text. In *Proceedings of the 27th International Conference on Computational Linguistics* (pp. 2104-2119).
- BRESSON, D., & DOBROVOL'SKIJ, D. (1995). Petite syntaxe de la « peur ». Application au français et à l'allemand. *Langue française*, 105, 107-119. DOI : [10.3406/lfr.1995.5297](https://doi.org/10.3406/lfr.1995.5297).
- CREISSEN, S., & BLANC, N. (2017). Quelle représentation des différentes facettes de la dimension émotionnelle d'une histoire entre l'âge de 6 et 10 ans? Apports d'une étude multimédia. *Psychologie française*, 62(3), 263-277. DOI : [10.1016/j.psfr.2015.07.006](https://doi.org/10.1016/j.psfr.2015.07.006).
- DIJKSTRA, K., ZWAAN, R. A., GRAESSER, A. C., & MAGLIANO, J. P. (1995). Character and reader emotions in literary texts. *Poetics*, 23(1-2), 139-157. DOI : [10.1016/0304-422X\(94\)00009-U](https://doi.org/10.1016/0304-422X(94)00009-U).
- DAVIDSON, D., LUO, Z., & BURDEN, M. J. (2001). Children's recall of emotional behaviours, emotional labels, and nonemotional behaviours: Does emotion enhance memory?. *Cognition & Emotion*, 15(1), 1-26. DOI : [10.1080/0269993004200105](https://doi.org/10.1080/0269993004200105).
- DAVIDSON, D. (2006). The role of basic, self-conscious and self-conscious evaluative emotions in children's memory and understanding of emotion. *Motivation and Emotion*, 30(3), 232-242. DOI : [10.1007/s11031-006-9037-6](https://doi.org/10.1007/s11031-006-9037-6).
- DIWERSY, S., GOOSSENS, V., GRUTSCHUS, A., KERN, B., KRAIF, O., MELNIKOVA, E., & NOVAKOVA, I. (2014). Traitement des lexies d'émotion dans les corpus et les applications d'EmoBase. *Corpus*, 13, 269-293.
- EKMAN, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6:3-4, 169–200. DOI : [10.1080/02699939208411068](https://doi.org/10.1080/02699939208411068).
- GROSS, M. (1995). Une grammaire locale de l'expression des sentiments. *Langue française*, 105, 70-87. DOI : [10.3406/lfr.1995.5294](https://doi.org/10.3406/lfr.1995.5294).
- HAMON, T., FRAISSE, A., PAROUBEK, P., ZWEIGENBAUM, P., and GROUIN, C. (2015). Analyse des émotions, sentiments et opinions exprimés dans les tweets : présentation et résultats de

- l'édition 2015 du défi fouille de texte (DEFT). Actes de la 22e conférence sur le Traitement Automatique des Langues Naturelles (TALN 2015), Jun 2015, Caen, France. HAL : [hal-01617180](https://hal.archives-ouvertes.fr/hal-01617180).
- LEWIS, M. (2008). Self-conscious emotions: embarrassment, pride, shame and guilt. In M. Lewis, J. M. Haviland-Jones, & L.F. Barrett (Eds.), *Handbook of emotions* (pp. 742–756). The Guilford Press.
- MATHIEU Y.Y. (2006) A Computational Semantic Lexicon of French Verbs of Emotion. In: SHANAHAN J.G., QU Y., WIEBE J., Édts., *Computing Attitude and Affect in Text: Theory and Applications*. The Information Retrieval Series, vol 20., chapitre 10, p. 109-124, Springer, Dordrecht. DOI : [10.1007/1-4020-4102-0_10](https://doi.org/10.1007/1-4020-4102-0_10).
- MICHELI R. (2010). L'émotion argumentée. L'abolition de la peine de mort dans le débat parlementaire français, Paris : Cerf.
- MICHELI, R. (2014). *Les émotions dans les discours : modèle d'analyse et perspectives empiriques*. De Boeck.
- MOHAMMAD, S. (2011). From once upon a time to happily ever after: Tracking emotions in novels and fairy tales. In *Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities* (pp. 105-114). Association for Computational Linguistics.
- PIOLAT, A., & BANNOUR, R. (2009). EMOTAIX: un scénario de Tropes pour l'identification automatisée du lexique émotionnel et affectif. *L'Année psychologique*, 109(4), 655-698. DOI : [10.4074/S0003503309004047](https://doi.org/10.4074/S0003503309004047).
- SAGOT (2010). The Lefff, a freely available and large-coverage morphological and syntactic lexicon for French. In *Proceedings of the 7th international conference on Language Resources and Evaluation* (LREC 2010). Istanbul, Turkey. HAL : [inria-00521242](https://hal.archives-ouvertes.fr/inria-00521242).
- WIDLÖCHER, A., & MATHET, Y. (2012, September). The glozz platform: A corpus annotation and mining tool. In *Proceedings of the 2012 ACM symposium on Document engineering* (pp. 171-180). HAL : [hal-01023774](https://hal.archives-ouvertes.fr/hal-01023774).
- ZWAAN, R. A., & RADVANSKY, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2), 162-85. DOI : [10.1037/0033-2909.123.2.162](https://doi.org/10.1037/0033-2909.123.2.162).