

11 mars 2021

Dimension émotionnelle des textes et compréhension

Delphine Battistelli

MoDyCo (UMR 7114, Paris Nanterre Université)

Séminaire laboratoire LIS (UMR 7020)

Equipe *TALEP*

Aix-Marseille Université



Plan

A. La compréhension de textes

A.1) en TAL

A.2) en psycholinguistique

B. La compréhension de textes par les enfants dans le cadre du projet TextToKids : premiers résultats

B.1) Prédire un âge recommandé

B.2) Annotation de la dimension émotionnelle

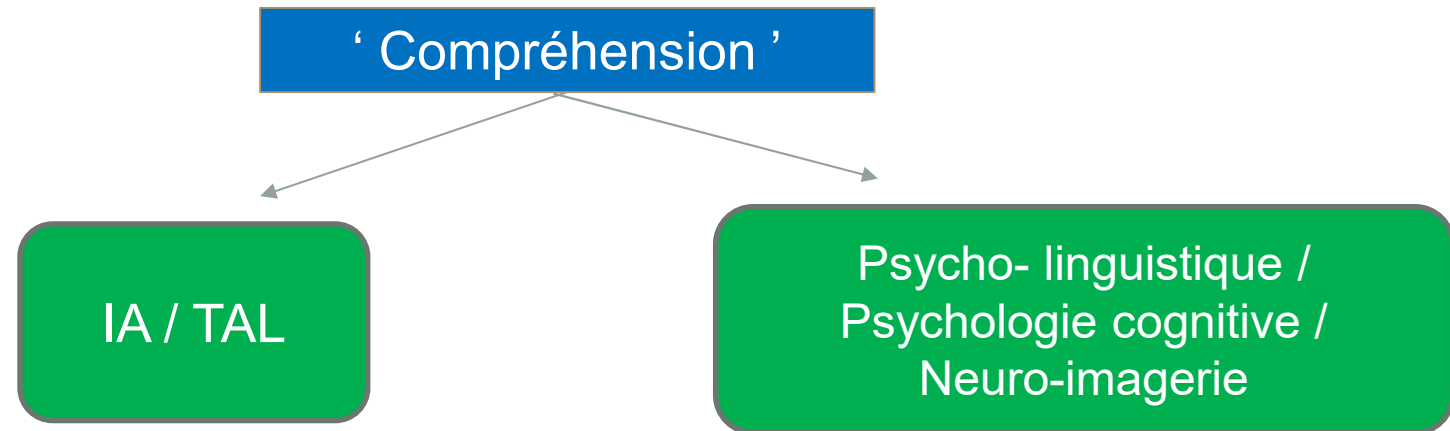
C. Explorer l'hypothèse de l'importance de la dimension émotionnelle pour la compréhension de textes

Plan

A. La compréhension de textes

A.1) en TAL

A.2) en psycholinguistique



Notions et méthodes mobilisées

- Représentation sémantique (ex. : Graphes conceptuels)
- Extraction d'information (ex. : les 4W)
- Systèmes de Q/R
- Schéma d'annotation (ex. : FrameNet)
- Compréhension Automatique de Textes (en anglais, *Machine Reading Comprehension*)

- Représentation mentale
- Dimension situationnelle
- Protocoles d'évaluation (ex. : tâche rappel, questions, EEG)

Du côté des ressources sémantiques

- Les dimensions sémantiques Temps, Espace et Cause sont unanimement mentionnées et explorées dans le champ de “la compréhension en TAL”
- La dimension Emotions *peut* l’être (cf. FrameNet pour l’anglais)

Du côté des ressources sémantiques

FrameNet, pour l'anglais (Baker et al., 1998)

- Modèle d'annotation sémantique pour l'anglais de situations prototypiques (appelées cadres) et de leurs participants (rôles)
- Comprend un lexique de déclencheurs, les lexèmes évoquant ces cadres et un ensemble d'annotations en cadres

-> contient 5 cadres relatifs au domaine *Emotion* : *Emotions*, *Emotions_by_possibility*, *Emotions_by_stimulus*, *Emotions_of_mental_activity*, *Emotions_success_or_failure*

- (Mohammad et al, 2014) : Corpus de tweets annoté avec cadre *Emotion* et 2 rôles investigués (parmi 12) : *Experiencer* (qui ressent), *Stimulus* (à propos de quoi)

FrameNet, pour le français

- **French FrameNet** (Candito et al, 2014 ; Djemaa, 2017) : Ressource pour le français créée à partir du FrameNet pour l'anglais, avec une approche centrée sur 4 domaines notionnels : *cause* (11 cadres), *communication langagière* (52 cadres), *position cognitive* (46 cadres), *transaction commerciale* (19 cadres)

-> ne contient pas de cadre *Emotion*

- **CALOR** (Béchet et al., 2017) : corpus de textes encyclopédiques annoté avec un ensemble de 54 cadres

-> ne contient pas de cadre *Emotion*

Du côté des ressources sémantiques en émotions

Lexique en émotions, pour l'anglais

- **NRC Emotion Lexicon** (Mohammad, 2011) : lexique constitué par un taliste. 8 catégories émotionnelles retenues : colère, dégoût, joie, peur, surprise, tristesse, confiance et anticipation/attente

Lexiques en émotions, pour le français

- **Emotaix** (Piolat & Bannour, 2009) : lexique. Constitué par des psychologues comportant 56 catégories organisées de manière hiérarchique (environ 2300 éléments)
- **Emolex** (Novakova, 2013) : lexique (noms, verbes, adjectifs). Constitué par des linguistes. Défini comme une « cartographie » linguistique des émotions. Pour 5 langues européennes (français, allemand, anglais, russe et espagnol). 9 champs sémantiques retenus : surprise, déception, respect, mépris, admiration, colère, jalousie, joie, tristesse

Corpus annoté en Emotions, pour l'anglais

- **SemEval 2019, tâche EmoContext** : émotions dans dialogues écrits entre un utilisateur et un agent conversationnel. 4 catégories émotionnelles retenues : joie, tristesse, colère, Autre

Corpus annoté en Emotions, pour le français

- **TALN 2015, Atelier DEFT** : identification de la classe de l'opinion, sentiment ou émotion parmi 18 classes : colère, peur, tristesse, dégoût, ennui, dérangement, déplaisir, surprise négative, apaisement, amour, plaisir, surprise positive, insatisfaction, satisfaction, accord, valorisation, désaccord, et dévalorisation

Corpus annoté en Emotions et Causes, pour l'anglais

- **Emotion Cause Extraction (ECE)** (Lee et al, 2010 ; Xia & Ding, 2019) : 4 catégories émotionnelles retenues : joie, tristesse, colère, peur, surprise. Extraction au niveau phrastique

Du côté des ressources sémantiques : bilan rapide

- Des ressources (lexiques, corpus annotés) existent pour les émotions
- Elles sont cependant beaucoup moins développées pour le français que pour l'anglais
- EN IA/TAL, elles ont été principalement conçues :
 - dans le champ de l'IHM ;
 - et, à la marge, dans le champ de l'analyse d'opinions/sentiments/émotions (devenu un champ à part entière en TAL depuis de nombreuses années)
- Le niveau de description linguistique des émotions reste dans tous les cas strictement lexical (cf. notion de déclencheur dans FrameNet)

Plan

A. La compréhension de textes

A.1) en TAL

A.2) en psycholinguistique

Compréhension et représentation de la situation sous-jacente au texte

- La plupart des modèles de compréhension proposés en psycholinguistique renvoient à des **dimensions sémantiques** mais qui ne sont pas nommées explicitement comme telles
- Elles sont nommées **dimensions situationnelles** (temps, espace, cause, personnages, ...)
- Ces modèles ont un point de vue commun sur la compréhension :
 - les performances des lecteurs en compréhension d'un texte ne peuvent s'expliquer que si ceux-ci ont construit une **représentation** de la **situation** (spatiale, temporelle, causale, ...) **sous-jacente au texte**
 - « la compréhension d'un texte est vue comme la construction d'une représentation situationnelle qui combine les propositions extraites du texte à celles qui ont été réactivées en mémoire à long terme afin de constituer **une représentation non du texte, mais de ce dont le texte parle.** » (Blanc, 2010)

Compréhension et représentation en mémoire *

- « **Comprendre** une histoire (lue ou entendue) implique que l'individu s'engage dans l'élaboration en **mémoire** d'une **représentation** des événements décrits.
- Cette représentation est supposée complexe et nécessite l'identification de la nature des relations situationnelles que les événements partagent, à savoir :
 - le cadre spatio-temporel dans lequel ils se produisent (càd., dimension **espace**, dimension **temps**),
 - les **personnages** qu'ils impliquent,
 - leur inscription sur la chaîne **causale** de l'histoire.
- La **représentation** que l'individu élabore en mémoire à partir de la situation décrite est ainsi envisagée comme étant « **multidimensionnelle** » ;
 - ce qu'attestent de nombreux travaux menés chez l'adulte (Zwaan et al., 1995; Zwaan & Radvansky, 1998; Boissery & Blanc, 2006; Tapiero & Blanc, 2001)
 - mais aussi quelques-uns réalisés chez l'enfant (cf. par ex. (Boisclair et al., 2004)) »

* Blanc N. (2010), La compréhension des contes entre 5 et 7 ans: Quelle représentation des informations émotionnelles ?, *Canadian Journal of Experimental Psychology*, 2010, Vol. 64, 4, 256–265

Consensus à propos de l'activité de compréhension

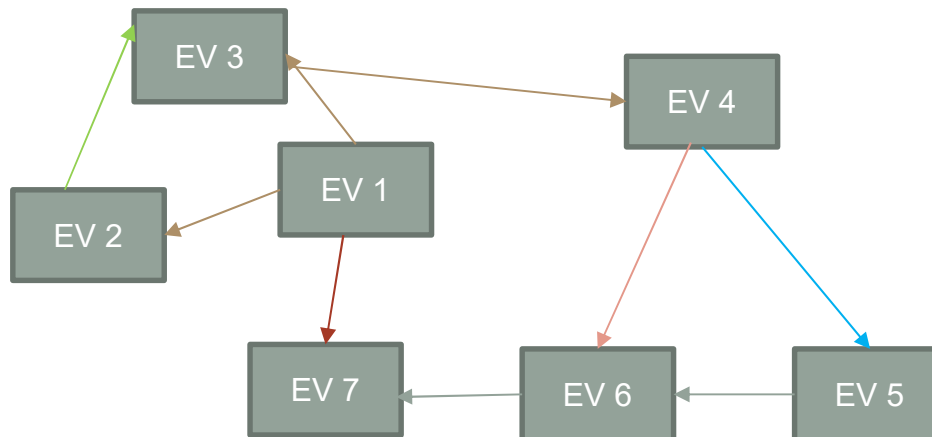
Compréhension ->

représentation de la **situation** sous-jacente au texte+ **représentation** que l'individu élabore en **mémoire**

Consensus à propos de l'activité de compréhension*

Est vue comme :

- un **réseau** d'évènements liés entre eux selon des **dimensions situationnelles** (Espace, Temps, Cause, Personnages, Emotion, ...)



- En rapport avec la mise en place d'**inférences** (de nature spatiale, temporelle, causale, émotionnelle, ...)

* (Blanc, 2021)

Exemple d'expérimentation chez des enfants de 5 à 7 ans*

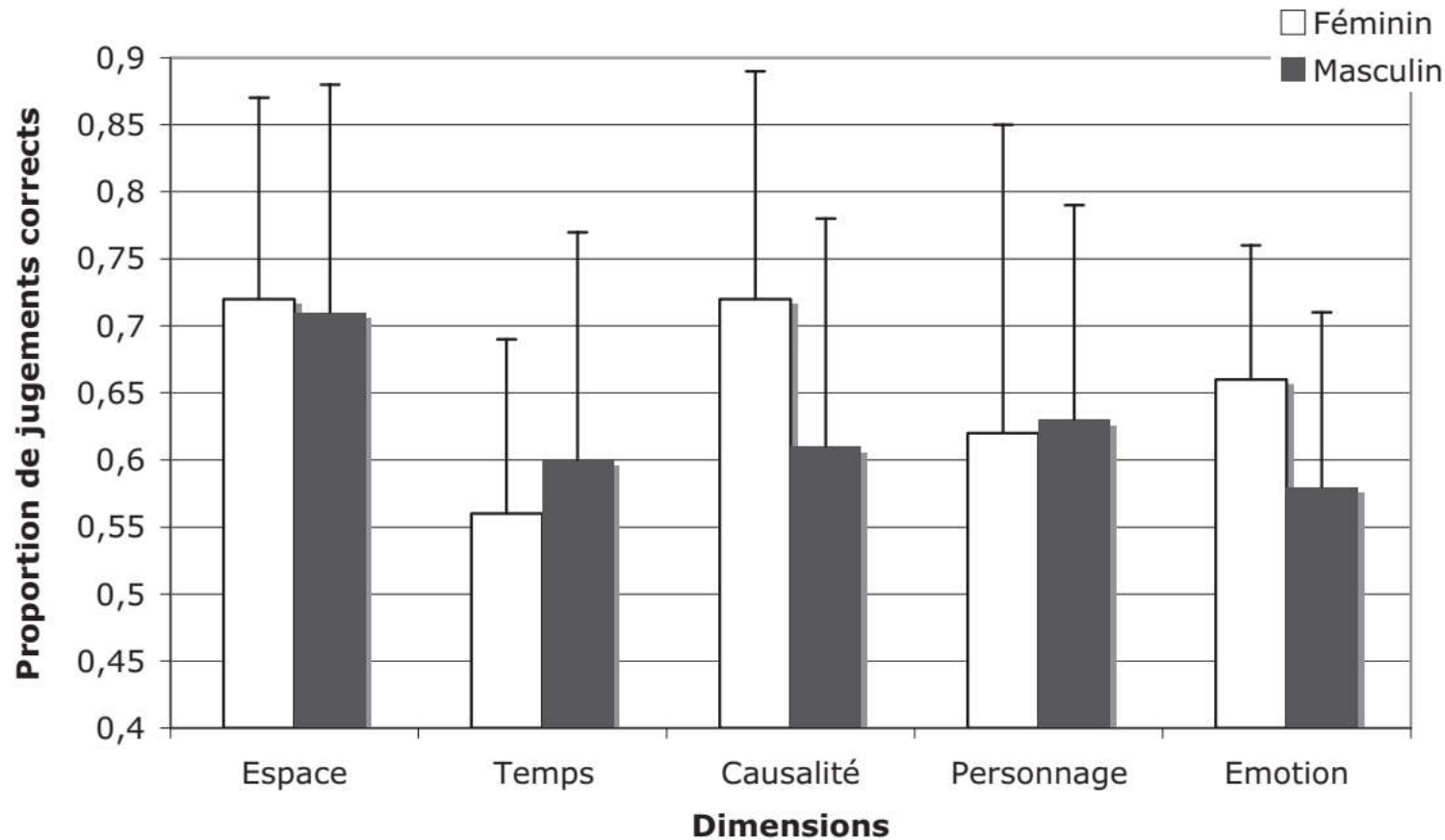


Fig. 1 : Tâches de compréhension de contes, enfants de 5 à 7 ans (moyennes des proportions de réponses correctes aux différentes dimensions situationnelles en fonction du sexe des enfants) – *(Blanc, 2010)

La dimension émotionnelle

- On retrouve ainsi dans ces modèles les dimensions sémantiques explorées dans le champ de la linguistique et du TAL
- Avec, en outre, une autre dimension situationnelle résolument investie : celle des émotions

La dimension émotionnelle – des émotions présentes dans un texte à leur rôle dans la mémorisation d'évènements

- Plusieurs travaux mettent ainsi en évidence le rôle majeur joué par la dimension émotionnelle d'un texte pour la compréhension du texte, tant chez les adultes (Dyer, 1983 ; Dijkstra et al., 1995) que chez les enfants (Davidson et al, 2001 ; Davidson, 2006)
- Les raisons principales en sont les suivantes :
 - Les émotions des personnages d'un texte facilitent la compréhension du texte car elles sont à l'origine d'**inférences** (ex. : *X est triste/embarrassé/fier... parce que Y a eu/aura/... lieu*).
 - Elles favorisent en outre la **mémorisation** de l'information, un **évènement associé à une émotion** ayant généralement une (ou des) conséquence(s) pour le personnage et/ou pour l'histoire alors que ce n'est pas forcément le cas d'un évènement qui n'est pas associé à une émotion
- Bagage lexical chez les enfants :
 - Chez les enfants, (Davidson et al, 2006) remarquent qu'une émotion « désignée » (i.e. via un label lexical, ex. : *triste/embarrassé/fier...*) dans le récit va être relativement peu restituée en tant que telle lors du rappel des évènements
 - La connaissance de l'émotion désignée (i.e. son label lexical) est d'ailleurs un pré-requis à l'observation de cette meilleure mémoire des évènements associés à une émotion

La dimension émotionnelle exprimée dans les textes

- Certaines des études menées sur les enfants, dont (Blanc, 2010 ; Blanc et Creissen, 2017), montrent en outre que, selon la manière dont elles sont exprimées, les émotions sont plus ou moins accessibles pour les plus jeunes
 - Il est démontré notamment que les émotions dites « comportementales » (ex. : X préfère s'éloigner de Y, X pleure) sont mieux comprises que les émotions « désignées » (ex. : X est embarrassé, X est triste) jusqu'à l'âge de 7 ans au moins
- La manière d'exprimer les émotions influence ainsi la capacité des enfants à comprendre le texte dans lequel elles se trouvent

Les émotions d'un texte, en fonction de la façon dont elles sont réalisées linguistiquement (on parlera aussi de leurs différent **modes d'expression**), constituent donc un enjeu pour la compréhension de textes par les enfants

Dimension émotionnelle et étapes développementales

- Certaines étapes développementales ont été identifiées concernant la compréhension des émotions elles-mêmes
- Davidson (2006) distingue trois **types** d'émotions (en se basant sur les travaux de (Ekman, 1992) et (Lewis, 2008) :

- **les émotions « de base »** (ou émotions primaires)

En nombre limité (colère, dégoût, joie, peur, surprise et tristesse) et relativement universelles (inter-individus et inter-cultures), elles sont maîtrisées dès la première année et sont associées à des expressions faciales stéréotypiques. Sur le plan de l'expression linguistique : acq. ≈ 3 ans

- **les émotions « de conscience de soi »** (ou « émotions sociales » (Blanc et Quenette, 2017)).

D'apparition un peu plus tardive (vers 1 an et demi), elles ne seraient pas associées aux mêmes situations et/ou aux mêmes réactions selon les cultures. Elles mettent en jeu des compétences cognitives plus complexes, notamment pour le traitement de l'information **sociale**. Elles nécessitent de focaliser son attention sur autrui (ex. : honte). Sur le plan de l'expression linguistique : acq. ≈ 6 ans

- **les émotions « de conscience de soi évaluatives »** (ou « sociales évaluatives » (Blanc et Quenette, 2017))

Rarement maîtrisées avant 3 ans, elles ont les mêmes caractéristiques que les émotions précédentes mais nécessitent de plus la capacité à se comparer ou à comparer son propre comportement à des **standards**, des **règles** ou des **objectifs culturels** internalisés. Ces éléments culturels font l'objet d'un apprentissage plus ou moins long chez l'enfant (ex. : fierté, culpabilité). Sur le plan de l'expression linguistique : acq. ≈ 9 ans

émotions « complexes »

La dimension émotionnelle exprimée dans les textes : 3 expériences chez les enfants

- Davidson (2006) :
 - Expé. de compréhension d'histoires menée sur enfants anglophones de 6, 8 et 10 ans confirme ces différentes étapes développementales
 - Enfants < 10 ans : expliquent beaucoup les émotions par des informations comportementales, surtout pour les émotions « de conscience de soi » ; quand ils sont confrontés à des émotions complexes (= de conscience de soi, évaluatives ou non), ils ne semblent pas en avoir une compréhension fine et ont tendance à les mentionner (voire peut-être les retenir) avec des termes liés aux émotions de base (ex. : *heureux* pour *fier*, *en colère* pour *gêné/embarassé*).
- (Baron-Cohen et al., 2010)
 - Expé. de compréhension de termes d'émotion menée sur enfants anglophones de 4 à 16 ans
 - Confirme aussi que les émotions de base sont accessibles plus tôt que les émotions complexes puisque les termes associés aux émotions de base sont beaucoup mieux compris et beaucoup plus tôt que ceux associés aux émotions complexes
- Blanc et Quenette (2017) :
 - Expé. de capacité à réaliser des inférences émotionnelles menée sur enfants francophones de 8 à 10 ans
 - Confirme ce dernier résultat. Les enfants retrouvent généralement la « coloration émotionnelle » mais ont plus de mal à saisir l'émotion « spécifique » suggérée par la situation, assimilant ainsi la fierté à la joie et la culpabilité à la colère

Une source potentielle de difficulté de compréhension pour les enfants est donc aussi le **type de l'émotion** :

- une émotion de base sera accessible très tôt
- tandis qu'une émotion de conscience de soi/sociale/complexe, nécessitant des compétences cognitives plus développées et des connaissances socio-culturelles plus importantes, ne sera pas maîtrisée avant 8-10 ans.

Protocoles expérimentaux auprès des enfants

- Temps lecture, questions, rappel ; pas d'EEG
- En France, accent mis surtout sur les difficultés des « faibles compreneurs » (anaphores, inférences, lexique ; ex. projet ANR Alector)
- Textes toujours de genre fictions/récits/... car textes familiers pour enfants dès leur plus jeune âge
 - Chez adultes : travaux très récents seulement en neuro-imagerie (EEG) avec prise en compte genre encyclopédique vs. journalistique (Aboud et al, 2019)
- Pas de test actuellement permettant d'évaluer les habiletés de compréhension de récits chez enfants francophones de l'école primaire (il en existe pour les mots et les phrases) ;
 - Sauf celui de (Maeder et Charlois, 2010) axé sur lexique et syntaxe et plutôt destiné à orthophonistes i.e. pas un test fondé sur modèles psycholinguistiques
 - Et celui de (Potocki et al., 2014) mais pas de dimension émotionnelle testée
- Matériau texte :
 - Compréhension = stade complexe car nécessite des habiletés à différents niveaux (lexical, syntaxique, sémantique, discursif) ;
 - Notamment au niveau des phénomènes de cohésion (moyens grammaticaux tq anaphores et connecteurs et lexicaux tq répétitions et synonymes) et de cohérence (sémantique et discursive)

L'hypothèse de l'importance de la dimension émotionnelle pour la compréhension de textes

- Du lien entre émotions et compréhension de textes dans l'ensemble des travaux en psycholinguistique, on retiendra la formulation/hypothèse sous-jacente suivante :

L'activation plus ou moins importante de la dimension émotionnelle dans un texte en favoriserait la compréhension dans la mesure où :

- i. elle améliorerait la **mémorisation** des *situations qui sont associées à des émotions* ;
étant entendu que la compréhension d'un texte peut être mesurée en relation avec la tâche de rappel notamment (càd la capacité à évoquer spontanément certaines situations décrites dans un texte) ou à l'aide de questions
- ii. elle participerait à la perception de la **cohérence** d'un texte ;
étant entendu que celle-ci peut être mesurée en relation avec la *capacité à mettre en relation des situations* évoquées dans un texte

- Offre une perspective méthodologique pour le TAL (schéma d'annotation)

Plan

A. La compréhension de textes

A.1) en TAL

A.2) en psycholinguistique

B. La compréhension de textes par les enfants dans le cadre du projet TextToKids : premiers résultats

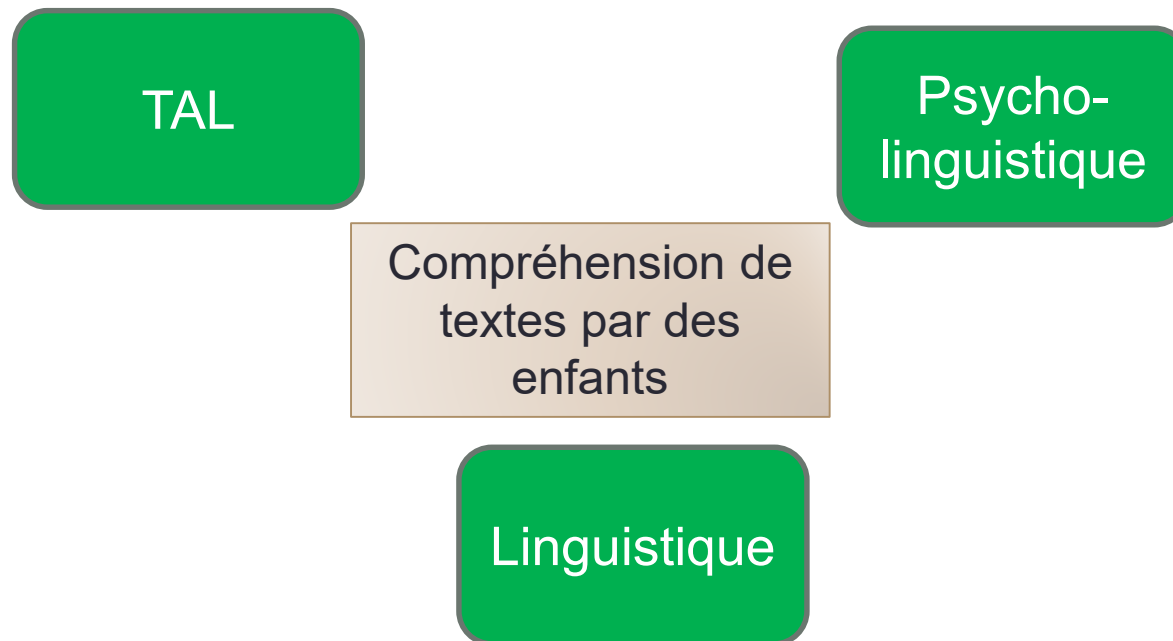
B.1) Prédire un âge recommandé pour un texte/une portion de texte

B.2) Annotation de la dimension émotionnelle

Le projet ANR TextToKids (2019-2023)

<https://texttokids.irisa.fr/>

- Scenario idéal : articuler entre eux des travaux existants dans les 3 disciplines pour l'étude du langage à destination des enfants jeunes lecteurs (7-12 ans)



Le projet ANR TextToKids (2019-2023)

<https://texttokids.irisa.fr/>

modyco

Synapse

UMR IRISA

Qwant

ALBERT
Petit journal illustré

LE PETIT
LIBÉ

Objectifs du projet

- étudier les **caractéristiques linguistiques et psycholinguistiques propices à une compréhension optimale d'un contenu informationnel (texte, portion de texte) par des enfants jeunes lecteurs (7-12 ans)**
- proposer des **briques logicielles de TAL qui intègrent ces caractéristiques**

Deux terrains expérimentaux privilégiés :

- *via* les journaux pour enfants Le Ptit Libé et Albert, qui ont pour visée de produire des articles décrivant des événements de l'actualité
- *via* le moteur de recherche dédié aux enfants Qwant Junior, qui indexe des textes de nature variée

Le projet ANR TextToKids (2019-2023)

Verrous scientifiques

- Deux problématiques centrales :

- ❑ Peut-on prédire l'adéquation d'un texte à une classe d'âge ?
- ❑ Comment traduire la notion de « complexité » d'un texte (ou d'une portion de texte) en une mesure objective, quantifiable à partir de descripteurs « élémentaires » ?

- Parmi les originalités du projet :

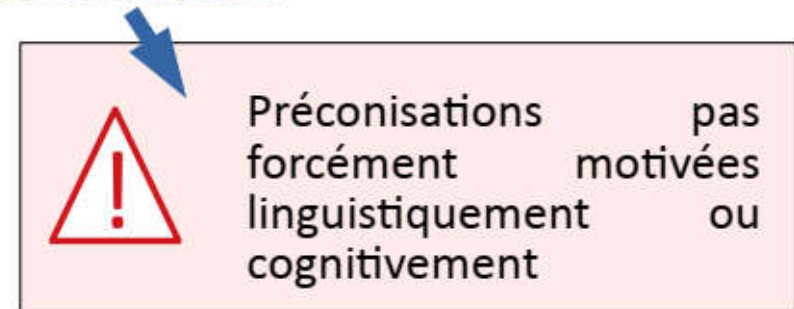
- Accent mis sur des descripteurs linguistiques **sémantiques** liés à l'expression des émotions (directement lexicalisées mais pas seulement ; ex. émotions suggérées par une situation), en lien avec d'autres dimensions sémantiques (par exemple, les dimensions de temps et de cause)
- Prise en compte de 3 genres textuels : journalistique, fictionnel et encyclopédique

Une première expérience menée en TAL : (Blandin et al., 2020)*

- *Peut-on prédire l'adéquation d'un texte ? → Peut-on prédire un âge recommandé pour un texte ?*
- Premier travail sur ce type de tâche à notre connaissance

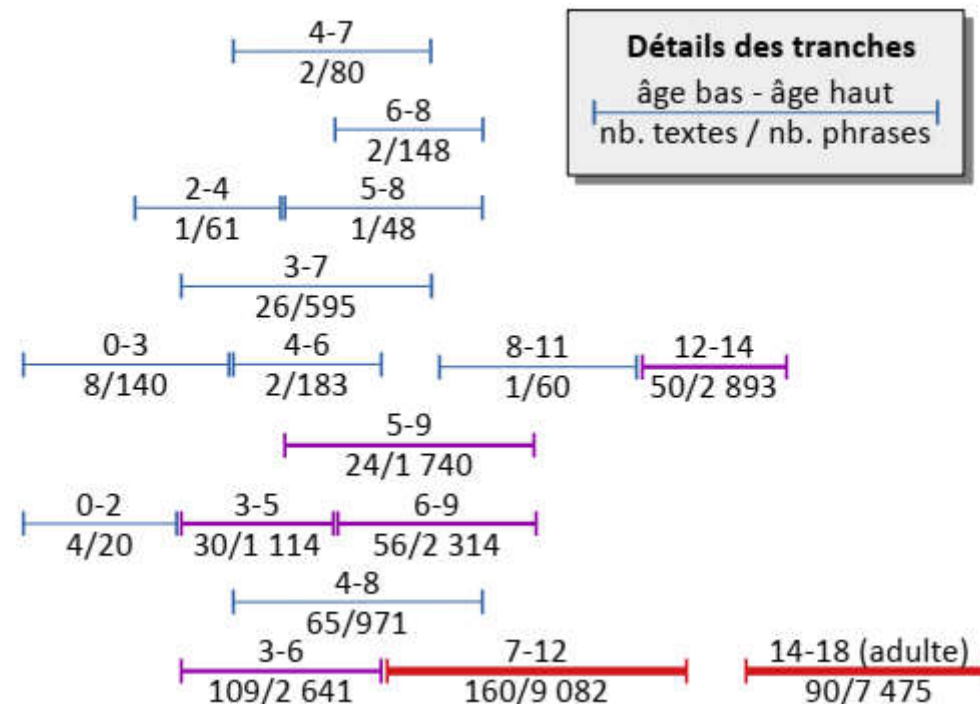
Une première expérience menée en TAL : (Blandin et al., 2020)*

- **Environ 600 textes, 30K phrases, 450K mots**
 - ~ 80 % enfant / 20 % adultes
 - Contes, romans, magazines, journaux
- **Annotations en tranches d'âge [X, Y] des auteurs et éditeurs**
 - Âge approximatif à partir duquel le texte peut être compris
 - **X** : âge bas , **Y** : âge haut
 - **(X+Y)/2** : âge moyen
- **Préparatifs**
 - Normalisation des textes, segmentation en phrases
 - Extraction et normalisation des features



Une première expérience menée en TAL : (Blandin et al., 2020)*

● Tranches d'âges diversifiées



Une première expérience menée en TAL : (Blandin et al., 2020)*

- Traitement comme une tâche de TAL classique
 1. Liste de descripteurs issus de la littérature SHS et plongements de mots
 2. Données annotées
 3. Différentes formalisations (régression, multitâche régression+classif., classif.puis régression)
 4. Comparaison avec des experts

Une première expérience menée en TAL : (Blandin et al., 2020)*

Résultats

- Prédiction d'un âge minimal recommandé
- Modèle performant pour cette tâche (erreur de $\pm 2,2$ ans)
- Plusieurs formalisations sans réelles différences
- Erreur de prédiction plus faible que celle d'experts

Perspectives

- Consolidation de la vérité terrain avec protocoles exp. auprès d'enfants au lieu préconisations éditeurs
- Prise en compte dimensions sémantiques, en particulier émotionnelle
- Augmentation et diversification du corpus

Plan

A. La compréhension de textes

A.1) en TAL

A.2) en psycholinguistique

B. La compréhension de textes par les enfants dans le cadre du projet TextToKids : premiers résultats

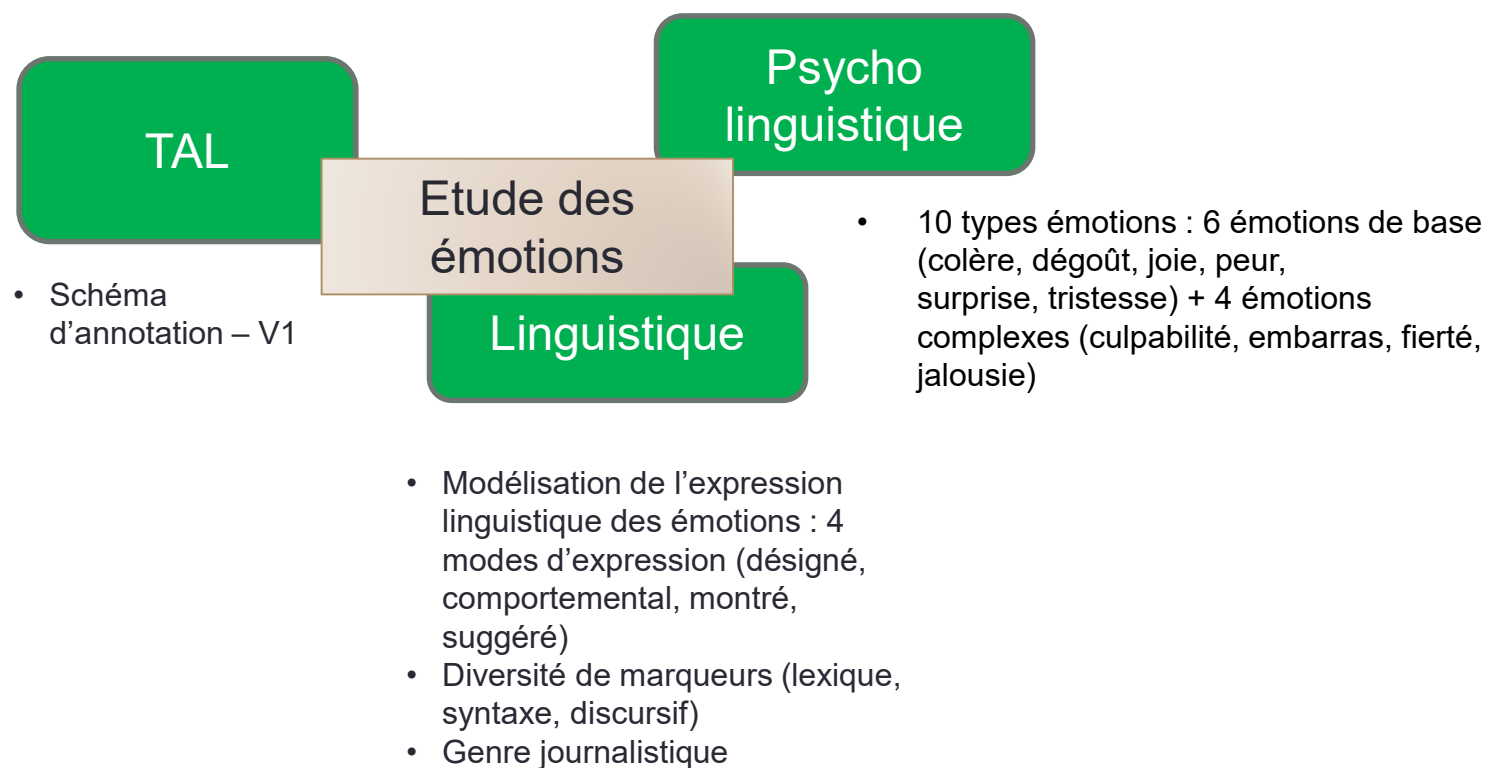
B.1) Prédire un âge recommandé

B.2) Annotation de la dimension émotionnelle

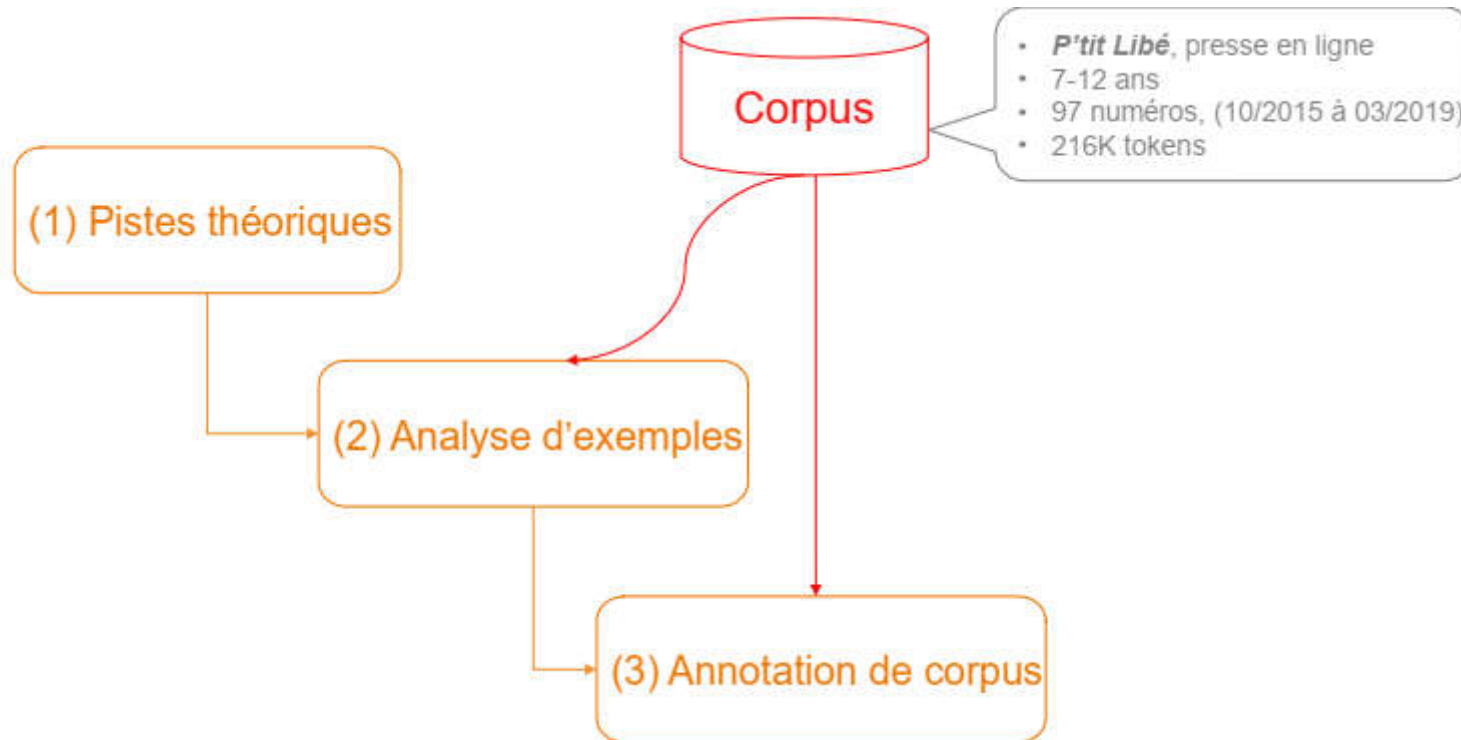
Une seconde expérience menée en TAL : (Etienne et al., 2020)*

- Annotation manuelle en émotions d'un corpus
- Plateforme d'annotation Glozz (Widlöcher et Mathet, 2012)

Approche tri-disciplinaire des émotions



Méthodologie



Modes d'expression linguistique des émotions

	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>

Exemples

Les 4 modes d'expression des émotions :

- « expression **comportementale** de l'émotion » : émotion exprimée par la description d'une manifestation physiologique (ex. pleurer, frissonner) ou comportementale (ex. taper du poing) de l'émotion ;

Ci-après en vert.

- émotion « **désignée** » : émotion désignée par le lexique émotionnel (ex. triste, effrayer) ;

Ci-après en jaune.

- émotion « **étayée** » : émotion inférée. L'inférence émotionnelle est étayée par la description d'une situation conventionnellement associée à une émotion (ex. des retrouvailles entre deux amis peuvent étayer la joie) ;

Ci-après en orange.

- émotion « **montrée** » : les caractéristiques de l'énoncé-même montrent que le locuteur était sous le coup d'une émotion lors de l'énonciation (ex. interjections, exclamation).

Ci-après en rouge.

Exemples annotés selon le schéma d'annotation

(a) En 2014, l'équipe du Brésil a été éliminée en demi-finale [...]. Cette défaite a rendu les Brésiliens tristes [...]. Ils pleuraient et déchiraient les tickets du match à la sortie du stade.

⇒ 'Tristesse' de l'entité « les Brésiliens » exprimée selon trois modes : comportemental, désigné et étayé.

(b) Grâce aux gaz à effet de serre, la température moyenne sur notre planète est de 15 °C. Sans eux, elle serait de... -18 °C, c'est très froid !

⇒ 'Surprise' de l'entité 'locuteur' exprimée selon un mode : montré.

(c) Au mois de décembre, elle a dû quitter sa maison à cause de la guerre.

⇒ 'Peur' de l'entité « elle » exprimée selon deux modes : comportemental et étayé.

Schéma d'annotation – V1

Catégories sémantiques : émotions et entités

Émotions [1], [3], [4]

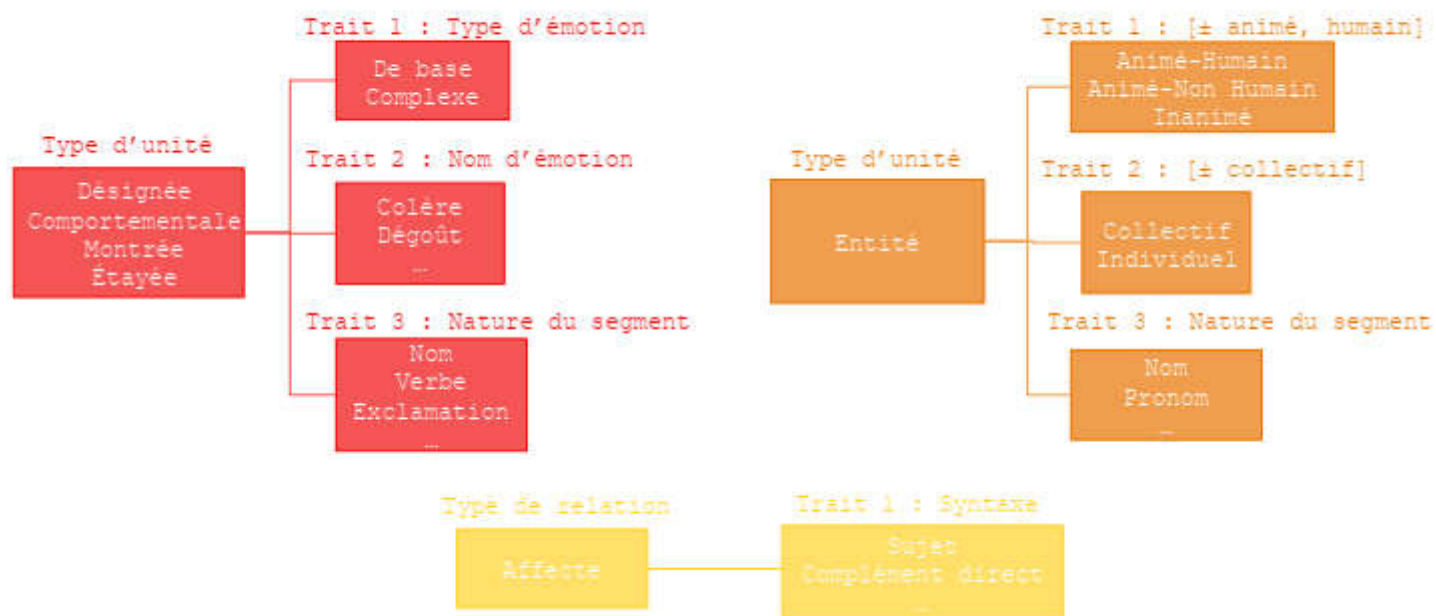
- 10 catégories émotionnelles
 - 6 de base
 - 4 complexes
- (+ Nature du segment)

Entité ressentant l'émotion [7], [8], [12]

- [± animé, humain]
 - [± collectif, individuel]
- (+ Nature du segment)

Schéma d'annotation – V1

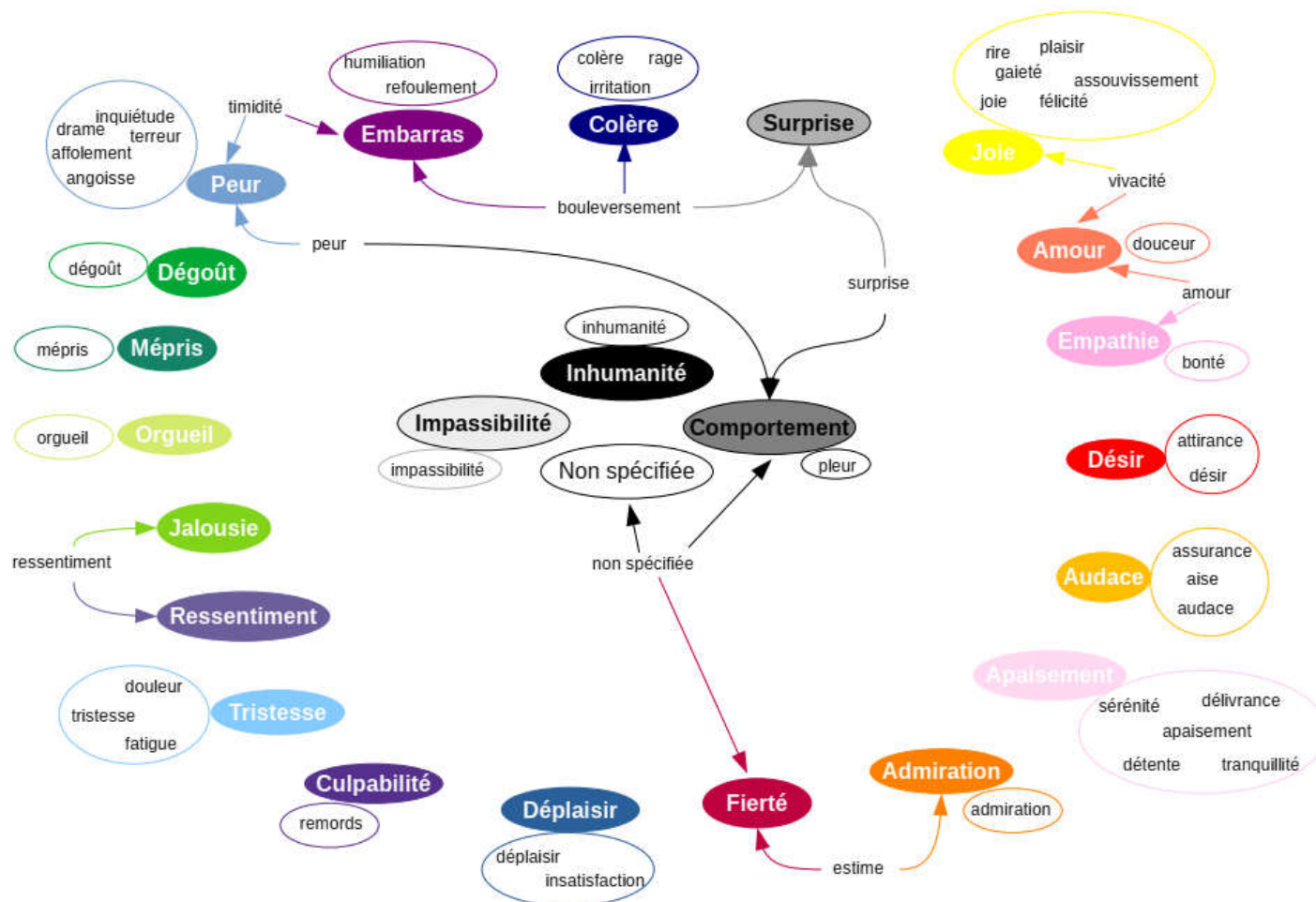
Schéma d'annotation



Ressources existantes mobilisées

Emotaix (Piolat et Bannour, 2009)

Leff (Sagot, 2010)

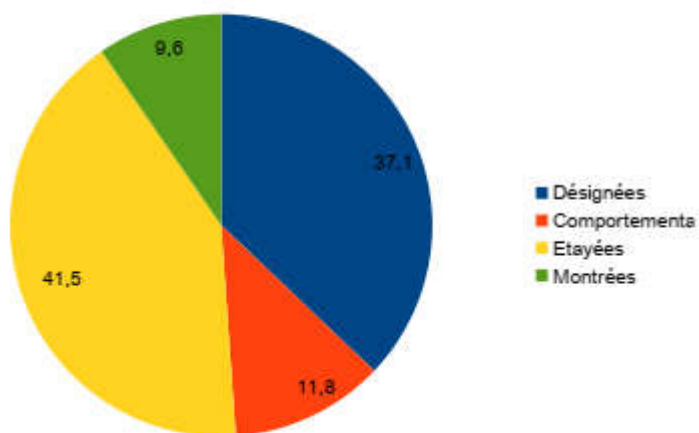




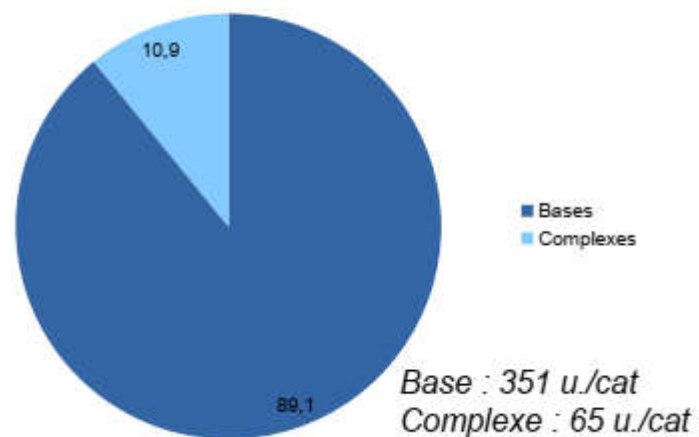
Observations sur le corpus

Total : 2368 unités émotionnelles

Proportions des modes d'expression (%)

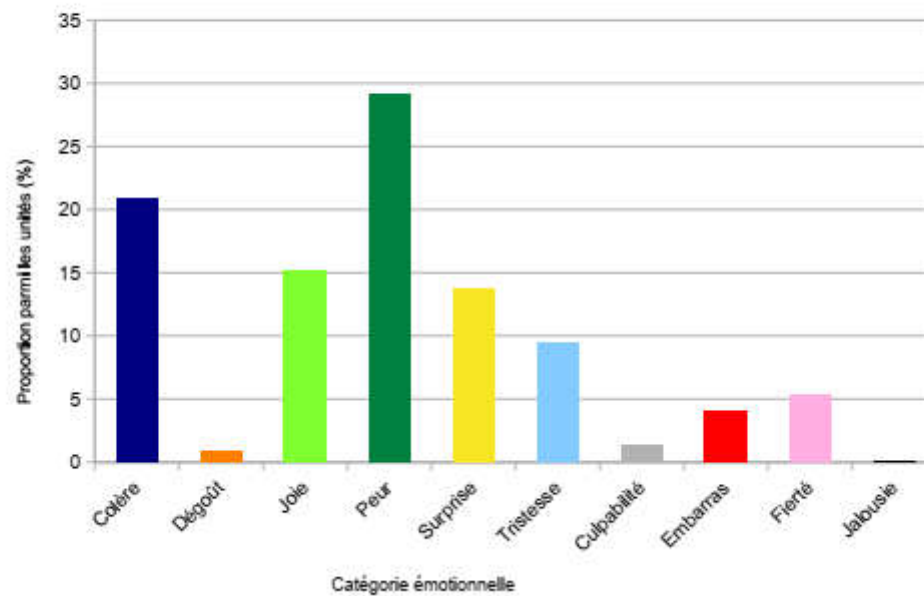


Proportions des types d'émotions (%)



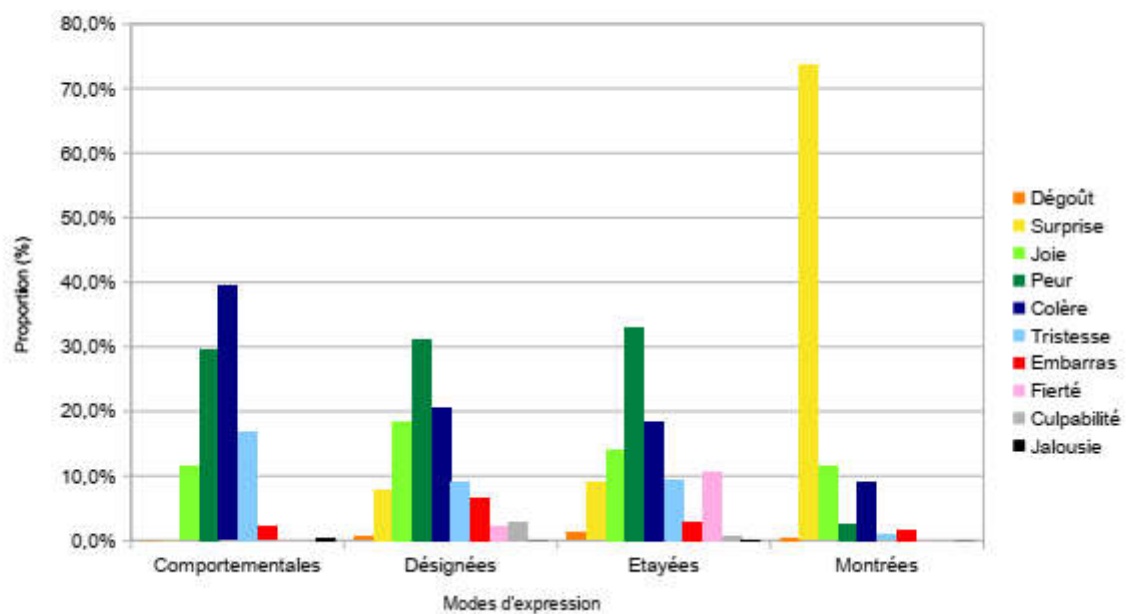
Observations sur le corpus

Proportions des catégories émotionnelles (%)



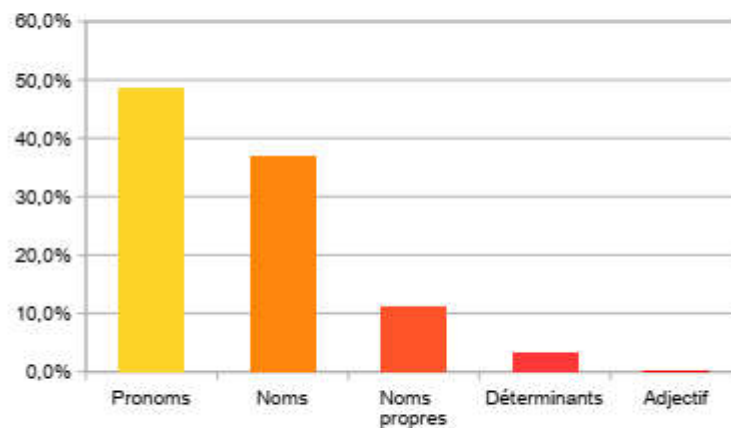
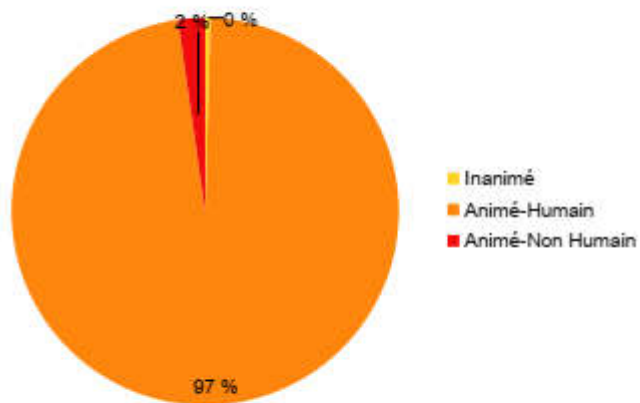
Observations sur le corpus

Proportions des catégories émotionnelles selon le mode d'expression (%)



Observations sur le corpus

Entités : 1110 unités au total



59,1 % [+ collectif]

Bilan

- Sur l'ensemble du corpus : majorité d'émotions négatives (peur et colère), cohérente avec les thèmes du P'tit Libé (thèmes de l'actualité).
- **Représentation des émotions variable selon le mode d'expression** : forte prédominance de la 'surprise' et peu de 'peur' dans les émotions montrées ; prépondérance de la 'colère' et quasi-absence de la 'surprise' dans les émotions comportementales.
- Ces observations quantitatives et l'analyse qualitative des annotations mettent en évidence :
 - ▶ **le recours important à l'émotion étayée**, catégorie complexe et hétéroclite ;
 - ▶ **l'association privilégiée entre des modes d'expression et certaines émotions** ;
 - ▶ **la pertinence des marqueurs linguistiques ne relevant pas du lexique émotionnel** pour la caractérisation des émotions ;
 - ▶ **des critères experts pour le repérage automatique des émotions** au sein d'une chaîne de traitement automatique des langues (TAL).

Limites

Limites et perspectives principales de notre tâche

Pas d'évaluation quantitative

- 1 seule annotatrice
- \nexists outils de calcul d'accord inter-annotateur (*unitizing*, relations)

Diversifier les types de textes annotés

- Genre textuel
- Auteur.e
- Public
- ...

Parmi les perspectives immédiates ouvertes

Linguistique

- Lien(s) avec la notion de prise en charge énonciative (Rabatel, 2013) et la temporalité (notion d'intensifieur, cf. Bordet, 2019) : *EN COURS*

Psycho linguistique

- Tests auprès de populations d'enfants : *EN COURS (fin prévue mai 2021)*

TAL

- Augmentation et diversification du corpus annoté : *REALISE*
- Schéma d'annotation-V2 et guide d'annotation : *EN COURS (fin prévue avril 2021)*

Plan

A. La compréhension de textes

A.1) en TAL

A.2) en psycholinguistique

B. La compréhension de textes par les enfants dans le cadre du projet TextToKids : premiers résultats

B.1) Prédire un âge recommandé

B.2) Annotation de la dimension émotionnelle

C. Explorer l'hypothèse de l'importance de la dimension émotionnelle pour la compréhension de textes

Rappel

L'hypothèse de l'importance de la dimension émotionnelle pour la compréhension de textes

- Du lien entre émotions et compréhension de textes dans l'ensemble des travaux en psycholinguistique, on retiendra la formulation/hypothèse sous-jacente suivante :

L'activation plus ou moins importante de la dimension émotionnelle dans un texte en favoriserait la compréhension dans la mesure où :

- i. elle améliorerait la **mémorisation** des *situations qui sont associées à des émotions* ;
étant entendu que la compréhension d'un texte peut être mesurée en relation avec la tâche de rappel notamment (càd la capacité à évoquer spontanément certaines situations décrites dans un texte) ou à l'aide de questions
- ii. elle participerait à la perception de la **cohérence** d'un texte ;
étant entendu que celle-ci peut être mesurée en relation avec la *capacité à mettre en relation des situations* évoquées dans un texte

- Offre une perspective méthodologique pour le TAL (schéma d'annotation)

Rendre possible l'exploration de cette hypothèse au travers d'un schéma d'annotation

i. mémorisation des **situations** qui sont **associées à des émotions** ;

=> Étude des situations associées à des émotions = cause et/ou conséquence d'une émotion

ii. perception de la **cohérence** d'un texte

=> Etude des phénomènes de structuration discursive, d'autant plus prégnants que prise en compte des 4 modes d'expression des émotions

(i) & (ii) -> **Cause – Emotion – Conséquence**

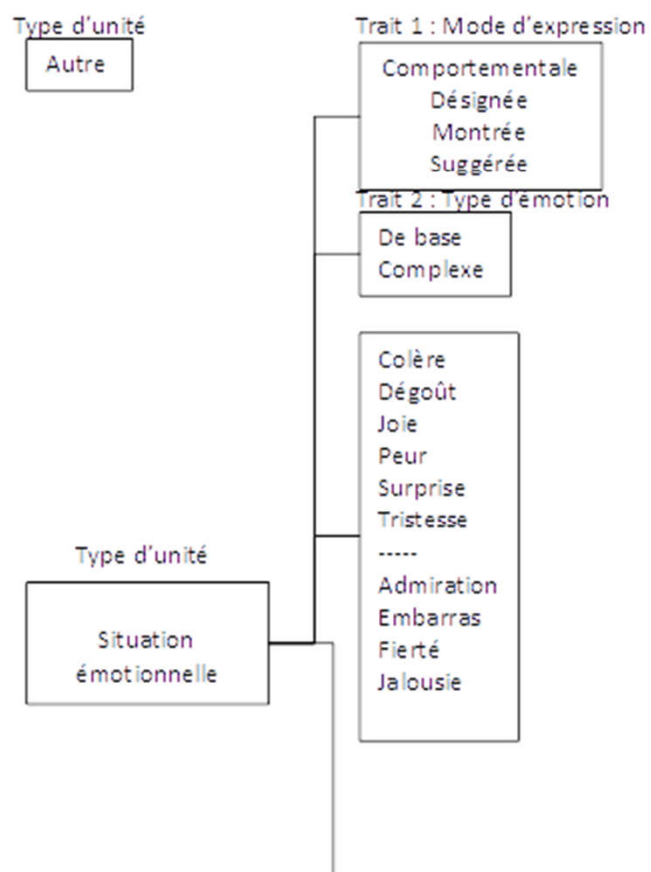
Sachant que ces 3 situations ont toutes des caractéristiques TAME

Dans notre schéma d'annotation, nous intégrons dorénavant :

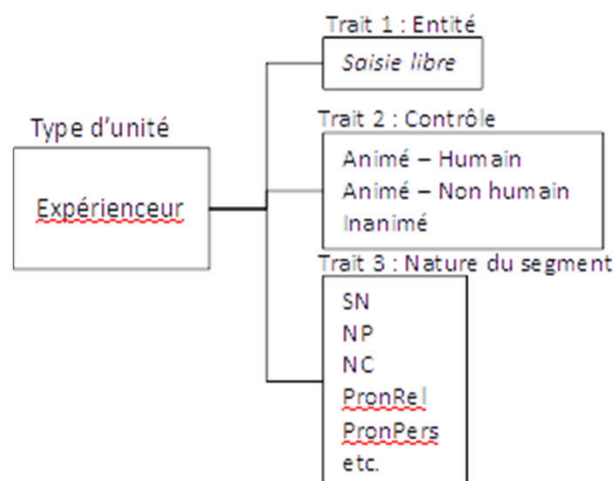
- **Les notions de CAUSE et de CONSEQUENCE**
- **Et celles de TAME**

Schéma d'annotation – V2

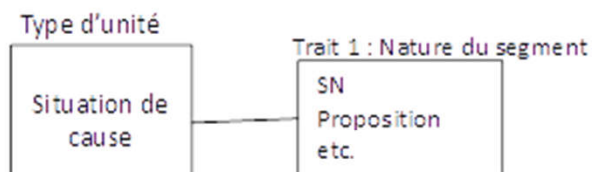
Unités émotionnelles :
Situation émotionnelle, Autre



Unité : Expérienceur



Unité : Situation de cause



Unité : Situation de conséquence

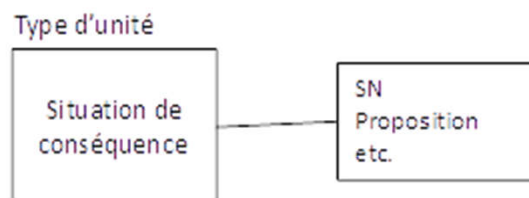


Schéma d'annotation – V2

Unités émotionnelles :
Situation émotionnelle, Autre

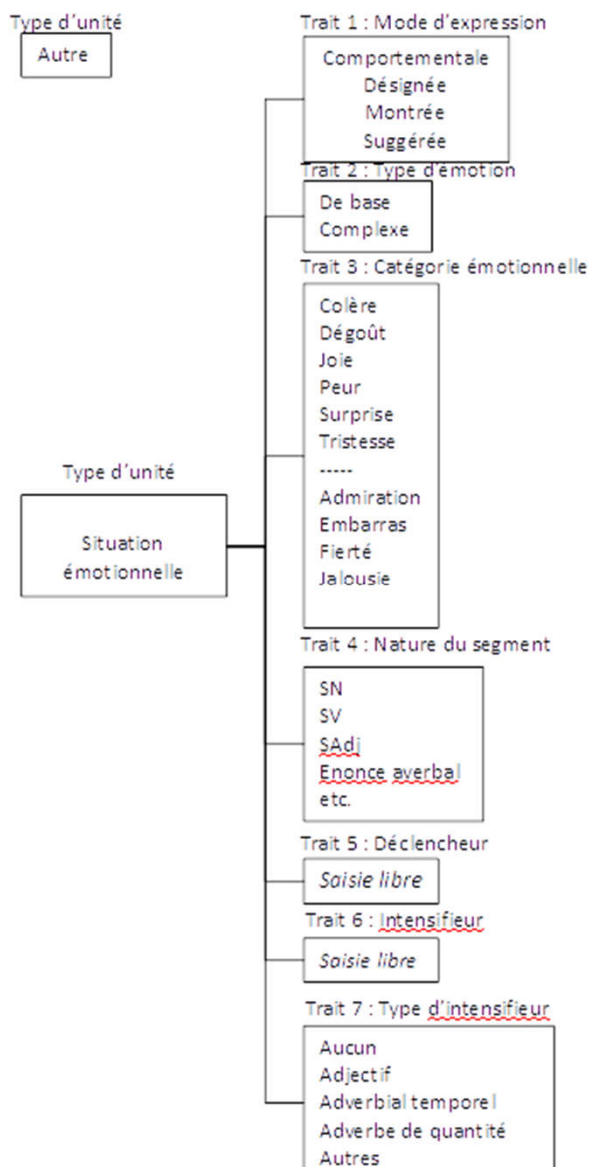


Schéma d'annotation – V2

Le schéma « Passage émotionnel » intègre des unités de type : « situation de cause », « situation émotionnelle », et « situation de conséquence »

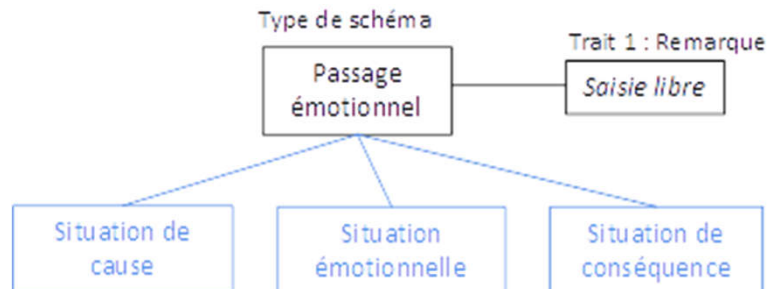


Schéma d'annotation – V2

Relation : Discontinue

Relation générale qui relie des unités discontinues

Type de relation

Discontinue

Relation : Affecte

*Relation qui relie une unité Situation émotionnelle à
une unité Expérienceur*

Trait 1 : Nature de la relation

Type de relation

Affecte

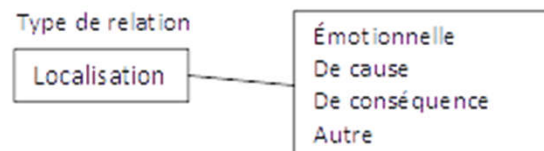
Sujet
Complément direct
etc.
Discours direct
Discours indirect
Contexte gauche
Contexte droit

Schéma d'annotation – V2

Relation : Localisation

Relation qui relie une unité temporelle (unité Situation ou unité Localisation temporelle) et une situation (unité Situation émotionnelle, Situation de cause, Situation de conséquence, Situation)

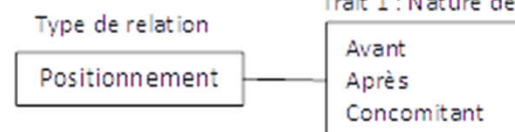
Trait 1 : Nature de la situation



Relation : Positionnement

Relation qui rend compte de la relation temporelle entre deux situations (unité Situation émotionnelle, Situation de cause, Situation de conséquence)

Trait 1 : Nature de la relation



Exemple (910-journ-Albert48-gilets-jaunes)

Albert - 21 décembre 2018

Depuis le 17 novembre, de nombreux Français **bloquent des routes** et **manifestent** contre le gouvernement et la politique du président Emmanuel Macron. Comme signe de leur **mécontentement**, ils ont **enfilé un gilet jaune fluo** : c'est pourquoi ce mouvement de **protestation** est appelé « mouvement des gilets jaunes ».

Qui sont les gilets jaunes et comment ce mouvement est-il apparu ?

À l'origine, le mouvement des gilets jaunes s'est formé suite à une **forte augmentation du prix** des carburants au mois de septembre. D'après les manifestants, l'État était en partie **responsable** de cette **hausse des prix** car il avait annoncé l'**augmentation des taxes** sur le carburant, en particulier sur le diesel, afin qu'il coûte à terme le même prix que l'essence.

Ainsi, le 17 novembre, des rassemblements de gilets jaunes se sont organisés un peu partout en France, notamment grâce aux réseaux sociaux. Des groupes se sont formés pour **bloquer des routes, des ronds-points ou des péages d'autoroute** durant toute la journée.

Puis de jour en jour, les mouvements de blocage ont continué, car les manifestants estimaient que le président Emmanuel Macron n'avait pas entendu leur **colère**.

- COLÈRE comportementale du personnage (“bloquent des routes”), (“manifestent”), (“enfilé un gilet jaune fluo”), (“protestation”), (“bloquer des routes, des ronds-points ou des péages d'autoroute”), (“démontraient”), (“revendications” ?), (“lutte”) + réclamer ??
- COLÈRE désignée du personnage (“mécontentement”), (“colère”)
- COLÈRE suggérée du personnage (“forte augmentation du prix”), (“responsable”), (“hausse des prix”), (“augmentation des taxes”), (“hausse des taxes”), (“ne gagnent pas suffisamment d'argent pour pouvoir vivre dignement”)

Exemple (910-journ-Albert48-gilets-jaunes)

Albert - 21 décembre 2018

Depuis le 17 novembre, de nombreux Français **bloquent des routes** et **manifestent** contre le **gouvernement et la politique du président Emmanuel Macron**. Comme signe de leur **mécontentement**, ils ont **enfilé un gilet jaune fluo** : c'est pourquoi ce **mouvement de protestation** est appelé « mouvement des gilets jaunes ».

Qui sont les gilets jaunes et comment ce mouvement est-il apparu ?

À l'origine, le mouvement des gilets jaunes s'est formé suite à une **forte augmentation du prix des carburants** au mois de septembre. D'après les manifestants, **l'État était en partie responsable** de cette **hausse des prix** car il avait annoncé **l'augmentation des taxes** sur le carburant, en particulier sur le diesel, afin qu'il coûte à terme le même prix que l'essence.

Ainsi, le 17 novembre, **des rassemblements de gilets jaunes** se sont organisés un peu partout en France, notamment grâce aux réseaux sociaux. **Des groupes se sont formés pour bloquer des routes, des ronds-points ou des péages d'autoroute** durant toute la journée.

Puis de jour en jour, les mouvements de blocage ont continué, car les manifestants estimaient que **le président Emmanuel Macron n'avait pas entendu leur colère**.

La cause de l'émotion est en **rose**.

- Remarque : certaines unités exprimant l'émotion sur le mode du suggéré semblent pouvoir correspondre à une cause de l'émotion sous-jacente

La conséquence de l'émotion est en **orange**.

- Remarque : certaines unités exprimant l'émotion sur le mode du comportemental semblent pouvoir correspondre à une conséquence de l'émotion sous-jacente

Exemple

(1112-fiction-Alector-Dragonne de minuit_ Laroche(diff))

Juché sur la cime d'un arbre mort, le corbeau contemplait pensivement la sombre bâtisse devant lui. Il songeait aux occupants des lieux. À Brunoît, le roi déchu, à Névé, sa femme, et à Mara, leur fille... Trois innocents condamnés à vivre enfermés ici, à Castelbrune, à tout jamais.

Dans un instant, il bouleverserait définitivement l'existence de l'un d'entre eux.

L'oiseau connaissait par cœur leur terrible histoire, elle lui avait été contée à maintes reprises.

Tout avait commencé quatorze ans plus tôt, alors qu'au palais le roi, passionné par l'étude des plantes médicinales, consultait un grimoire dans son laboratoire.

Nour, sa sœur aînée, avait surgi, affolée :

– Viens vite, Angus est brûlant de fièvre et ne peut se lever.

Nour, veuve depuis fort longtemps, chérissait son fils au plus haut point, redoutant sans cesse qu'il lui arrive malheur .

Exemple

(1112-fiction-Alector-Dragonne de minuit_ Laroche(diff))

Juché sur la cime d'un arbre mort, le corbeau contemplait **pensivement** la sombre bâtisse devant lui. Il songeait aux occupants des lieux. À Brunoît, le roi déchu, à Névé, sa femme, et à Mara, leur fille... Trois **innocents condamnés** à vivre enfermés ici, à Castelbrune, **à tout jamais**.

Dans un instant, il bouleverserait définitivement l'existence de l'un d'entre eux.

L'oiseau connaissait par cœur leur **terrible** histoire, elle lui avait été contée à maintes reprises.

Tout avait commencé quatorze ans plus tôt, alors qu'au palais le roi, passionné par l'étude des plantes médicinales, consultait un grimoire dans son laboratoire.

Nour, sa sœur aînée, **avait surgi**, **affolée** :

– **Viens vite**, Angus est brûlant de fièvre et ne peut se lever.

Nour, veuve depuis fort longtemps, chérissait son fils au plus haut point, **redoutant** sans cesse qu'il lui arrive **malheur** .

- Corbeau / narrateur
 - TRISTESSE comportementale (“pensivement”) ?
 - TRISTESSE suggérée (“innocents condamnés”) Intensifieur (temporel) : “à tout jamais”
 - TRISTESSE / PEUR désignée (“terrible”)
- Nour
 - PEUR comportementale (“avait surgi”)
 - PEUR désignée (“affolée”), (“redoutant”)
 - PEUR suggérée (“brûlant de fièvre”), (“malheur”)
 - PEUR montrée (“viens vite”)

Bilan et développements actuels

- Un travail en phase initiale encore (projet TTK légèrement décalé en lien avec crise Covid et tests enfants non réalisés)
- Chaîne TAL de (Blandin et al, 2020) en cours de développement chez Synapse avec travaux en cours sur émotions et temporalité à intégrer ensuite
- Annotation manuelle corpus fictions et encyclopédique avec schéma d'annotation V2
- Définition d'un protocole d'évaluation de compréhension de textes en écoles primaires (Marseille et Montpellier) :
 - textes « naturels » diversifiés en 3 genres (3 par genre, classes CE2 –CM2) ;
 - équilibrage des questions vis-à-vis de 2 critères : émotions (implicites/explicites) et temporalité au sens strict (ordonnancement et ancrage)
 - Résultats attendus : 1) dimension émotionnelle mieux suivie que dimension temporelle en CE2 et de façon identique en CM2 ;2) dimension ancrage temporel mieux suivie en CM2 qu'en CE2

- Corpus actuel :

	Train				Valid.				Test			
Genre	Texts	Sent.	Age Range	Mean	Texts	Sent.	Age Range	Mean	Texts	Sent.	Age Range	Mean
Encyclop.	738	21,960	11.33-15.71	13.52	201	6,495	10.62-15.13	12.87	135	4,325	11.52-15.86	13.69
Newspaper	793	18,245	09.44-13.85	11.64	147	3,224	09.38-13.78	11.58	189	4,118	09.49-13.86	11.67
Fiction	966	47,277	09.06-11.85	10.45	255	12,471	07.86-10.90	09.38	231	13,186	08.66-11.83	10.24
Overall	2,509	87,882	09.73-13.26	11.49	605	22,261	08.90-12.57	10.74	559	21,701	09.41-13.03	11.22

Table 1: Summary of the age prediction train, valid. and test sets

Bibliographie

- About K.S., Bailey S.K., Del Tufo S.N., Barquero L.A., Cutting L.E. (2019), Fairy Tales versus Facts: Genre Matters to the Developing Brain. *Cereb Cortex*. 2019 Dec 17;29(11):4877-4888.
- Baker C.F., Fillmore C.J., Lowe J.B. (1998), The Berkeley FrameNet Project, in Proc. ACL'1998, Montreal, Quebec, Canada.
- Baron-Cohen S., Golan O., Wheelwright S., Granader, Y. (2010), Emotion word comprehension from 4 to 16 years old: a developmental survey. *Frontiers in evolutionary neuroscience*, 2, 109.
- Béchet F., Damnati G., Heinecke J., Marzinotto G., Nasr A. (2017), CALOR-Frame : un corpus de textes encyclopédiques annoté en cadres sémantiques. In ACor4French – Les corpus annotés du français - Atelier TALN, Orléans, France.
- Blanc N. (2010), Lecture et habiletés de compréhension chez l'enfant. Dunod.
- Blanc N., Quenette G. (2017). La production d'inférences émotionnelles entre 8 et 10 ans: quelle méthodologie pour quels résultats?. *Enfance*, (4), 503-511.
- 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.
- Blanc N. (2021) : Les émotions dans la compréhension de récits: Actualité de la recherche chez l'enfant, Présentation Séminaire Psycholinguistique du laboratoire MODYCO, 9 février 2021.
- Blandin A., Lecorvé G., Battistelli D., Étienne A. (2020), Recommandation d'âge pour des textes, in Actes JEP-TALN'2020.
- Boisclair A., Makdissi H., Sanchez C., Fortier C., Sirois P. (2004), La structuration causale du récit chez le jeune enfant. In Actes du 9eme colloque de IAIRDF.
- Boissery A., Blanc N. (2006), Les dimensions situationnelles dans la compréhension de textes scientifiques, historiques et journalistiques: Vers une extension du modèle d'indexage d'événements. *Cahiers Romains de Sciences Cognitives*.
- Bordet L. (2019), Le rôle des marqueurs d'intensification dans l'expression des émotions en anglais, *Lexis [En ligne]*, 13 | 2019.
- Candito M., Amsili P., Barque L., Benamara F., Chalendar G., Djemaa M., Haas P., Huyghe R., Mathieu Y., Muller P., Sagot B., Vieu, L. (2014), Developing a French FrameNet : Methodology and First results. In Proc. LREC'2014, Reykjavik, Iceland.

Bibliographie

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*, n° 62, 2017, p. 263-277.

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.

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.

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.

Djemaa M. (2017), Stratégie domaine par domaine pour la création d'un FrameNet du français : annotations en corpus de cadres et rôles sémantiques, PhD, Université Paris Diderot.

Dyer M. G. (1983), The role of affect in narratives. *Cognitive Science*, 7(3), 211-242.

Étienne A., Battistelli D., Lecorvé G. (2020), L'expression des émotions dans les textes pour enfants : constitution d'un corpus annoté, in Actes JEP-TALN'2020.

Ekman P. (1992), An argument for basic emotions. *Cognition and Emotion*, 6:3-4, 169–200.

Lee S., Chen Y., and Huang C-R. (2010), A text-driven rule-based system for emotion cause detection. In Proc. NAACL HLT' 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text, pages 45–53.

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.

Maeder C., Charlois A. L. (2010), Validation d'un test de compréhension écrite de récit : « La forme noire ». *Glossa*, 108, 69–85

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 Proc ACL-HLT'2011, Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities (pp. 105-114).

Mohammad S., Zhu X., Martin J. (2014), Semantic Role Labeling of Emotions in Tweets, in Proc. 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media, Baltimore, Maryland

Bibliographie

Novakova I. (2013), EMOLEX Le lexique des émotions dans cinq langues européennes : sémantique, syntaxe et dimension discursive : Rapport final projet ANR DFG Emolex (2009-2013).

Piolat A. et Bannour R. (2009), EMOTAIX: un scénario de Tropes pour l'identification automatisée du lexique émotionnel et affectif. *L'Année psychologique*, 109(4), 655-698.

Potocki A., Bouchafa H., Magnan A., Ecalle J. (2014), Évaluation de la compréhension écrite de récits chez l'enfant de 7 à 10 ans : vers des profils de compreneurs, *Revue européenne de psychologie appliquée* 64 (2014) 229–239.

Rabatel A. (2013), Ecrire les émotions en mode emphatique. *Semen. Revue de sémio-linguistique des textes et discours*, (35).

Rahman R., Lecorvé G., Etienne A., Béchet N., Chevelu J., Battistelli D. (2020), Mama/Papa, Is this Text for Me?, in *Actes COLING'2020*.

Sagot B. (2010), The Lefff, a freely available and large-coverage morphological and syntactic lexicon for French. In *Proc LREC'2010, Istanbul, Turkey*.

Tapiero I., N. Blanc (2001), Vers la prise en compte de la caractéristique multidimensionnelle des représentations mentales construites à partir des textes narratifs : apports théoriques, empiriques et questions, *L'Année psychologique*, 2001 101-4, pp. 655-682.

Widlöcher A., Mathet, Y. (2012), The Glozz Platform: A Corpus Annotation and Mining Tool, In *Proc. 2012 ACM Symposium on Document Engineering*, pp.171-180.

Xia R., Ding Z. (2019), Emotion-Cause Pair Extraction:A New Task to Emotion Analysis in Texts, in *ProcACL'2019*, pages 1003–1012, Florence, Italy.

Zwaan R. A., Radvansky G. A. (1998), Situation models in language comprehension and memory. *Psychological bulletin*, 123(2), 162.