Un baromètre affectif effectif

Yves Bestgen, Cédrick Fairon, Laurent Kevers

Un baromètre affectif effectif[1]

Yves Bestgen[1], Cédrick Fairon[2], Laurent Kevers[2]

Université catholique de Louvain

(1) FNRS, Faculté de psychologie, Unité de psychologie sociale et des organisations

(2) Faculté de philosophie et lettres, Centre de traitement automatique du langage

Place Cardinal Mercier - 1348 Louvain-la-Neuve - Belgique

e-mail :yves.bestgen@psp.ucl.ac.be; fairon@tedm.ucl.ac.be ; kevers@tedm.ucl.ac.be

Abstract

The work objective of the research project presented here is to develop an Information Extraction method for selecting sentences containing named entities (person names or organisation names) and for rating automatically their affective valence (pleasant-unpleasant). In order to achieve this goal, a named entities extraction system is used together with a lexical analyser based on electronic dictionaries containing affectiveness scores. In parallel, we have built a reference corpus which permits to evaluate the scoring system by comparing its results to the judgment of human subjects

Résumé

L'objectif de la recherche rapportée ici est de développer une technique d'extraction d'information permettant de déterminer automatiquement la valence affective de phrases qui mentionnent des noms de personnalités ou de sociétés. Pour ce faire un extracteur d'entités nommées est associé à un programme d'analyse lexicale faisant appel à des dictionnaires de valence affective. Un corpus de référence est établi pour mesurer les performances du système proposé en les comparant à des jugements humains.

Mots-clés : valence affective, extraction d'information, entités nommées

1. Introduction

L'objectif de la recherche rapportée ici est de développer une technique d'extraction d'information permettant de déterminer automatiquement la valence affective de phrases qui mentionnent des noms de personnalités ou de sociétés. A terme, il s'agit de pouvoir répondre à des questions comme : *le nom de tel homme politique ou de telle dirigeante d'entreprise est-il plus souvent mentionné dans un contexte linguistique plutôt positif, agréable ou bien à l'inverse dans un contexte plutôt négatif, désagréable ?* Combinée à un instrument qui extrait des journaux accessibles en ligne les passages d'articles qui mentionnent le nom d'une personne, une telle technique pourrait servir de base pour construire un « baromètre affectif ». Elle permettrait en effet de suivre au jour le jour la valence affective des phrases dans lesquelles on fait référence à une personne donnée et d'en étudier l'évolution dans le temps.

A un niveau plus général, les questions qui portent sur la manière dont un élément est présenté et évalué dans un texte (comme de décider si la critique d'un film ou un commentaire boursier est positif ou négatif) sont particulièrement complexes pour les techniques d'extraction d'information (Wilks, 1997 ; Das et Chen, 2001 ; Pang et al., 2002). Elles constituent donc un

[1] Nous remercions pour leur aide Bernadette Dehottay et Sophie Piérard ainsi que les 10 étudiants qui ont contribué à la réalisation de cette recherche.

domaine de test particulièrement intéressant pour confronter différentes approches, mais aussi pour déterminer la manière la plus efficace de les combiner. Bien plus, cette complexité contraste avec la simplicité de l'approche la plus fréquemment employée pour répondre à ces questions : se baser exclusivement sur les mots qui composent un texte pour en déterminer la valence affective.

Initialement, cette approche a été développée dans le champ de l'analyse de contenu. Déjà en 1965, Heise a proposé de constituer un dictionnaire de norme d'évaluation en demandant à des juges d'évaluer sur la dimension agréable-désagréable un échantillon des mots[2] les plus fréquents d'une langue. Depuis lors, des dictionnaires pour différentes langues ont été constitués (Heise, 1965 ; Hogenraad et al., 1995 ; Whissell et al., 1986). Ces dictionnaires ont été employés pour évaluer la valence affective de textes, mais aussi d'unités plus petites comme des phrases (Bestgen, 1994). La procédure proposée par Heise (1965) est très simple. Dans un premier temps, on dresse la liste des mots différents et de leur fréquence dans l'unité textuelle. Cette liste est comparée à un dictionnaire qui contient un ensemble de mots dont on connaît la valence affective. A chaque fois qu'un mot se trouve dans les deux listes, on affecte la valeur indiquée dans le dictionnaire au mot du texte. Enfin, on calcule la moyenne des valeurs connues. Malgré le caractère rudimentaire de cette technique, des arguments en faveur de sa validité ont pu être apportés (Anderson et al., 1982 ; Bestgen, 1994 ; Whissell et al., 1986). Plus récemment, des chercheurs en extraction d'information ont adapté cette procédure afin de la simplifier en n'utilisant que des mots fortement négatifs et des mots fortement positifs (Das et Chen, 2001; Turney et Littman, 2002). Dans cette version, la valence affective d'un texte est déterminée en soustrayant le nombre de mots négatifs du nombre de mots positifs contenus dans ce texte.

Le caractère simpliste de cette approche basée sur les mots considérés individuellement a fait l'objet de plusieurs critiques (Bestgen, 1994 ; Pang et al., 2002 ; Polanyi et Zaenen, 2003) qui plaident pour la combinaison d'informations lexicales et d'analyses linguistiques plus complexes. Par exemple, Polanyi et Zaenen (2003) soulignent la nécessité de prendre en compte les négations, mais aussi certains connecteurs (*Altough Boris is brillant in math, he is a horrific teacher*) et les opérateurs modaux (*If Mary were a terrible person, she would be mean to her dogs*). Il faut toutefois noter que les arguments empiriques à la base de ces critiques reposent sur l'analyse de quelques exemples le plus souvent construits à dessein par les chercheurs. Dépasser cette approche intuitive est nécessaire si l'on veut pouvoir évaluer les gains qu'apporte la prise en compte de chacun des dispositifs linguistiques censés affecter la valence d'une phrase.

La présente étude s'inscrit dans cette direction. Elle vise deux objectifs : tout d'abord, constituer un corpus de phrases dont la valence affective est connue. Dans ce but, nous avons sélectionné d'une manière largement aléatoire (voir ci-dessous) 702 phrases dans un corpus journalistique. Ces phrases ont été présentées à 10 juges dont la tâche était de les évaluer sur la dimension agréable-désagréable (section 3). Comme l'indique la section 3.2.1, l'accord inter-juges obtenu est très élevé. Ce corpus pourra donc être utilisé pour tester les différentes techniques proposées pour prédire la valence affective de phrases. Il pourra aussi être employé d'une manière plus heuristique afin de mettre en lumière les facteurs linguistiques qui modulent l'intensité affective perçue par des juges. Le second objectif de notre étude est d'évaluer sur ce corpus l'efficacité d'une approche basée exclusivement sur les mots qui composent les phrases à évaluer (section 4).

[2] classes ouvertes!

2. Constitution du corpus

1) Le matériel de départ est composé de l'ensemble des articles du quotidien national belge *Le Soir* parus entre début janvier et fin avril 1995. Une édition relativement ancienne du journal a été choisie pour éviter une trop grande sensibilité des juges par rapport à des sujets d'actualité.

2) Un système d'extraction (décrit dans Fairon et Watrin, 2002) a ensuite été utilisé pour sélectionner toutes les phrases contenant un ou plusieurs noms de personne et analyser ces séquences (c'est-à-dire reconnaître le prénom, le nom et les éventuelles informations du type profession, âge, nationalité, titre, etc.). Il s'agit d'un système basé sur des dictionnaires électroniques et des transducteurs et faisant appel au logiciel de traitement de corpus Unitex[3]. Des quelque 66500 phrases identifiées dans les textes par cette procédure, nous avons sélectionné celles d'une longueur supérieure à 8 mots et contenant au moins un nom propre apparaissant 10 fois ou plus dans l'ensemble des phrases. La justification de cette étape est que pratiquement il est peu intéressant d'étudier la valence affective associée à des personnes trop rarement mentionnées dans la presse. Au terme de cette seconde sélection, il restait un peu plus de 11000 phrases.

3) Chacune des phrases a été modifiée automatiquement de manière à remplacer le nom et la description de fonction par un prénom générique de genre adéquat (Marie, Jean, Pierre, etc.). Ceci pour éviter que les juges ne soient influencés par l'image *a priori* positive ou négative qu'ils peuvent avoir des personnes évoquées dans la presse.

4) Deux échantillons ont été extraits de ce corpus.

   a) On a extrait aléatoirement un échantillon de 700 phrases. Lors de cette étape, on a pris en compte la longueur en mots des phrases de façon a obtenir un échantillon représentatif pour cette variable. Cet échantillon est supposé représentatif du type de phrases incluant des noms propres que l'on peut trouver dans ce genre de journal. La procédure de sélection utilisée ne garantit toutefois pas qu'il contienne un nombre suffisant de phrases très agréables ou très désagréables pour pouvoir ultérieurement tester des techniques d'évaluation automatique. Pour cette raison, un deuxième échantillon de phrases a été constitué.

   b) Un deuxième échantillon de 371 phrases a été extrait du corpus et présenté à deux juges indépendants dont la tâche était d'indiquer dans quelle mesure le contenu de chaque phrase évoquait pour eux une idée plutôt désagréable, neutre ou agréable. Ils disposaient pour ce faire d'une échelle à 5 points allant de très désagréable à très agréable en passant par neutre. Sur la base de leur jugement, on a sélectionné toutes les phrases qui étaient de même polarité pour les deux juges et qu'un des deux juges au moins avait évaluées comme très désagréable ou comme très agréable.l

5) Enfin, l'ensemble du matériel a été passé en revue par deux personnes afin d'en éliminer toutes les phrases incomplètes, syntaxiquement incorrectes ou incompréhensibles hors de leur contexte original.

Le corpus final est composé de 702 phrases dont 606 ont été sélectionnées par la procédure aléatoire et 96 par la procédure de pré-jugements.

[3] http://www-igm.univ-mlv.fr/~unitex/

3. Évaluation du corpus

Dix étudiants de la faculté de philosophie et lettres de l'UCL ont été recrutés, par voie d'affiche, pour participer à cette étude. Une compensation financière était offerte en échange de leur participation afin de les inciter à effectuer la tâche avec toute l'attention nécessaire.

3.1 Modalités

Dans les consignes, nous expliquions aux participants que nous souhaitions constituer un corpus de phrases dont l'intensité affective (agréable - désagréable) est connue. Leur tâche était de lire une série de phrases publiées dans la presse et d'indiquer si, selon eux, le contenu de chacune d'elles évoquait une idée plutôt désagréable, neutre ou agréable. Pour donner leur avis, ils disposaient de l'échelle suivante :

				neutre			
Très désagréable	**-3**	**-2**	**-1**	**0**	**+1**	**+2**	**+3**	Très agréable

Les sept échelons recevaient une dénomination verbale basée sur les mots : très (désagréable ou agréable), moyennement, un peu et neutre.

Les 702 phrases du corpus étaient présentées dans un ordre aléatoire différent pour chaque participant. La difficulté majeure de ce genre de tâche réside dans le manque initial d'ancrage de l'échelle (Bestgen, 1994). Les participants lisant les phrases une à une, il leur est difficile de situer les premières phrases par rapport aux pôles extrêmes. Afin de pallier cette difficulté, douze phrases ont été ajoutées au matériel et présentées à tous les participants dans un ordre unique au début de la tâche. Ces phrases avaient été sélectionnées dans le matériel pré-testé de sorte à exprimer l'ensemble du continuum d'évaluation depuis le pôle très désagréable jusqu'au pôle très agréable.

Concrètement, chaque participant répondait au questionnaire au moyen d'un ordinateur. Les phrases étaient successivement affichées sur l'écran juste au-dessus de l'échelle. Le participant répondait en cliquant sur le bouton correspondant à son évaluation. Un bouton « valider » lui permettait de confirmer son choix et de faire s'afficher la phrase suivante. A tout moment, il pouvait interrompre la tâche. Les consignes précisaient qu'il devait effectuer une pause d'au moins une heure vers le milieu de la tâche. En moyenne, les participants ont mis 15 secondes pour juger chaque phrase.

3.2 Résultats

Ce travail vise deux objectifs principaux : constituer un corpus de phrases dont l'intensité affective (agréable - désagréable) est connue et tester une technique d'estimation de la valence affective de phrases basée exclusivement sur les mots qui les composent. Ces deux objectifs ne peuvent être atteints que si un accord inter-juges suffisamment élevé est observé dans la tâche d'évaluation. Les premières analyses vérifient cette condition. Dans un deuxième temps, nous comparerons les évaluations moyennes faites par les juges aux valeurs qui peuvent être prédites sur la base de la valence affective des mots qui composent les phrases.

3.2.1 Accord inter-juges

Pour déterminer l'accord inter-juges, nous avons calculé le coefficient alpha de Cronbach. Il s'élève à 0.93. Deux indices complémentaires ont été calculés. La corrélation moyenne entre les réponses d'un participant et la moyenne de tous les autres participants est de 0.75. La

corrélation moyenne entre les réponses fournies par deux participants est de 0.60. Ces valeurs signalent un accord inter-juges très élevés.

3.2.2 Analyses des évaluations

La figure 1 présente la distribution des valeurs moyennes de valence affective pour les deux échantillons séparément. Comme le montre la distribution pour le second échantillon, la procédure de pré-jugement a bien permis d'accroître dans le corpus la proportion de phrases à forte valence. Dans l'échantillon aléatoire, on observe un large étalement des valeurs entre -2 (moyennement désagréable) et +2 (moyennement agréable).

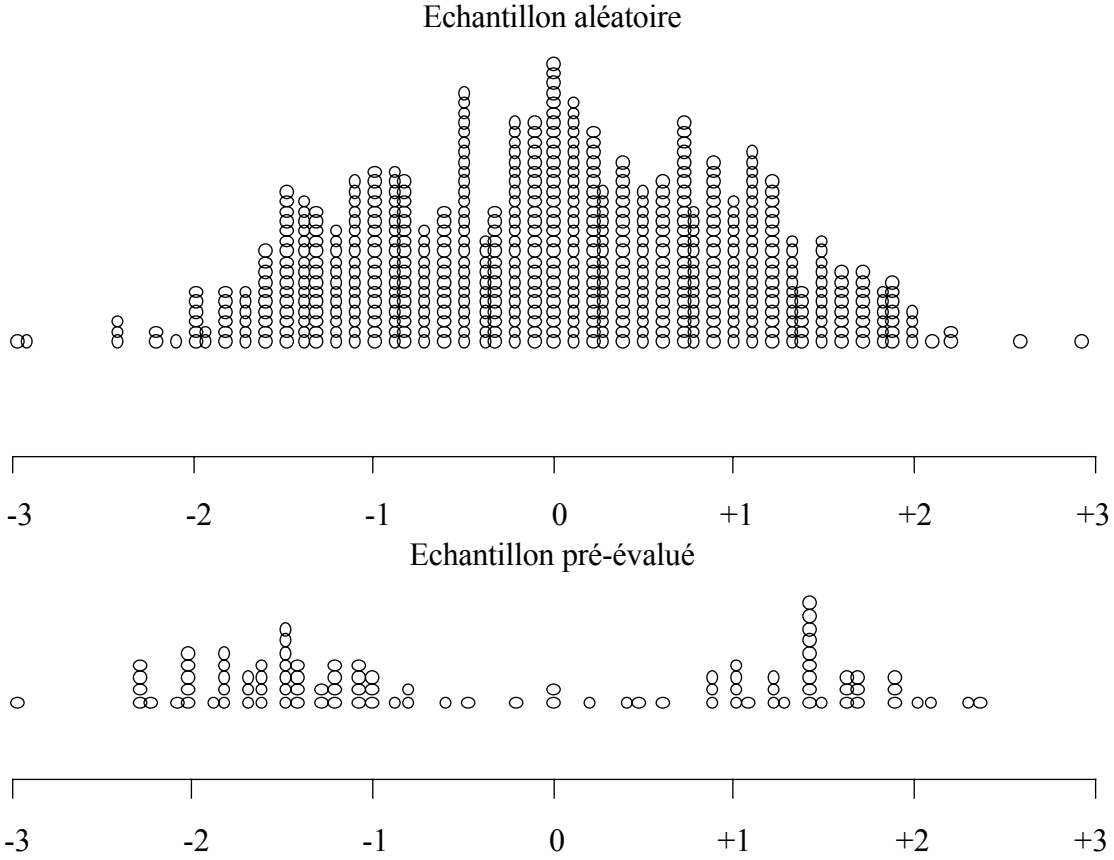

Figure 1 *: Distribution des valeurs moyennes de valence affective pour les deux échantillons.*

4. Estimation de la valence affective

4.1 Relation entre les valeurs prédites et les valeurs réelles

Pour déterminer l'intensité affective des phrases sur la base des mots qui les composent deux dictionnaires ont été employés.

Le dictionnaire générique est composé de 3289 mots qui ont été évalués par un minimum de 15 juges sur une échelle à 7 points allant de très désagréable à très agréable (Hogenraad et Bestgen, 1989; Bestgen, 1994, 2002). Le tableau 1 donne quelques exemples de mots du dictionnaire sélectionnés aléatoirement à différents niveaux d'intensité affective.

Mot	Valence	Mot	Valence
détresse	1.4	contrôlable	3.5
imbécile	1.4	outil	4.3
tristesse	1.6	risquer	4.5
hostilité	2.2	entier	4.9
impassible	2.6	revenir	5.0
superstitieux	2.8	admiratif	5.7
hâte	3.1	doux	6.0
ambigu	3.2	sincérité	6.1

Tableau 1 *: Valences affectives sur une échelle allant de très désagréable (1) à très agréable (7).*

La limitation principale de ce premier dictionnaire est qu'il est construit sur la base d'un échantillon de mots constitué une fois pour toutes, et donc qu'il n'est pas spécifiquement adapté au lexique employé dans le corpus analysé. Or, Bestgen (1994) a montré qu'on prédisait d'autant mieux l'intensité affective d'une unité textuelle qu'on prenait en compte l'orientation affective d'un plus grand nombre des mots qui la composent. Afin de vérifier l'impact de cette limitation, un second dictionnaire a été constitué en demandant à deux juges de parcourir la liste de tous les mots présents dans le corpus et d'en extraire ceux qui évoquent une idée agréable et ceux qui évoquent une idée désagréable. Sur la base de cette liste, l'intensité affective d'une phrase est calculée en soustrayant le nombre de mots désagréables du nombre de mots agréables contenus dans cette phrase.

Ces deux listes ont été employées pour dériver deux scores de valence affective appelés *générique* et *spécifique* selon le dictionnaire utilisé. Ces analyses ont été effectuées sur les phrases lemmatisées au moyen du programme « TreeTagger » de Schmid (1994)[4]. Seuls les mots catégorisés comme des noms, des verbes, des adjectifs ou des adverbes ont été pris en compte.

Afin d'estimer l'efficacité des deux scores lexicaux, nous les avons comparés aux valences affectives estimées par les juges au moyen du coefficient de corrélation de Pearson. Cette corrélation est de 0.39 pour le dictionnaire générique et de 0.56 pour le dictionnaire spécifique, valeurs toutes deux significatives au seuil de 0.0001. Ces corrélations ne sont pas modifiées lorsqu'on prend en compte la longueur des phrases ; la corrélation partielle pour le dictionnaire générique est égale à 0.37 et celle pour le dictionnaire spécifique à 0.56.

Si ces résultats sont encourageants, ils sont loin d'être parfaits. La figure 2 présente un diagramme de dispersion avec les évaluations moyennes des juges en ordonnée et les valeurs prédites par le dictionnaire spécifique en abscisse. La taille des cercles traduit le nombre de phrases qui présentent un couple de valeurs donné. Ce graphique montre qu'une partie importante de l'imprécision résulte du grand nombre de phrases qui reçoivent un score proche de zéro sur l'indice lexical. Ce constat est confirmé par les analyses suivantes. Lorsqu'on ne prend en compte que les 452 phrases dont l'indice spécifique est différent de 0 la corrélation passe à 0.62. Elle passe à 0.75 pour les 182 phrases dont l'indice spécifique est inférieur à -1 ou supérieur à 1 et à 0.83 pour les 75 phrases dont l'indice lexical est supérieur à -2 et +2. Il est à noter que cet effet s'observe tant pour l'échantillon aléatoire que pour l'échantillon pré-évalué.

[4] Disponible à l'adresse http://www.ims.uni-stuttgart.de/projekte/corplex/DecisionTreeTagger.html ainsi qu'au Cental, sous forme de service en ligne (adresse http://cental.fltr.ucl.ac.be/outils.html).

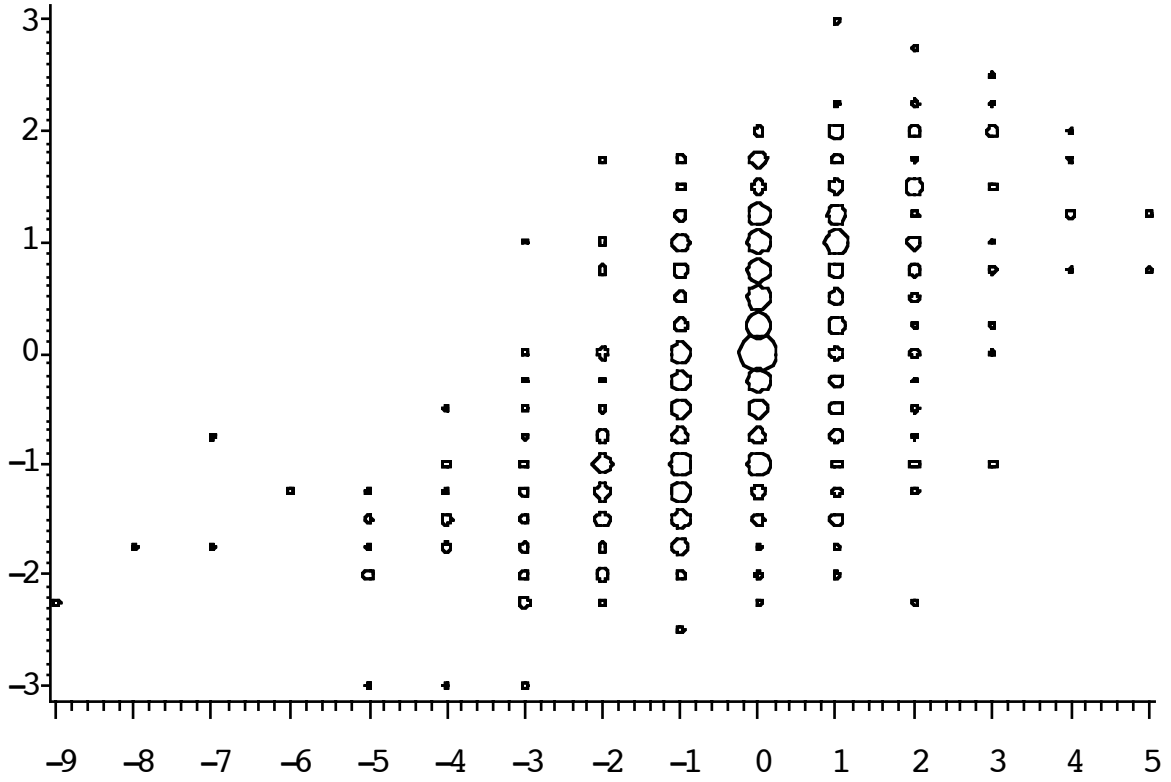

Figure 2 : ***Diagramme de dispersion avec les évaluations moyennes des juges en ordonnée et les valeurs prédites par le dictionnaire spécifique en abscisse.***

Intuitivement, on peut deviner que si une valeur très positive est obtenue par le programme pour une phrase que les juges ont évaluée très négativement, c'est que des phénomènes linguistiques échappant à la simple mesure lexicale entrent en jeu. Pour en savoir plus, nous avons passé en revue les phrases dont la valeur prédite était très différente de la valeur estimée par les juges. Nous présentons ici succinctement quelques-unes des observations que nous avons pu réaliser.

– La présence d'un verbe ou expression verbale modifie (inverse) la valeur affective de certains mots ou syntagmes : *s'en prendre à, affecter de, singer, se moquer de, hypothéquer, mettre en péril, mettre en cause*, etc. Par exemple, dans la phrase suivante, extraite de notre corpus, on constate la présence d'un certain nombre d'items lexicaux positifs (*fraternité, solidarité, liberté,* etc.) qui expliquent la prédiction positive donnée par le programme. Si l'avis des juges sur cette phrase est clairement négatif, c'est en raison du verbe principal de la phrase qui lui donne une valeur inverse à celle suggérée par la présence de ces mots positifs :

En frappant Tom, Rome s'en prend au-delà des clivages philosophiques, à tous les hommes et à toutes les femmes qui placent au premier rang les valeurs de fraternité, de solidarité, de liberté et de l'esprit.

— De la même manière, on peut facilement imaginer une situation opposée où un verbe inverse le caractère négatif de certains mots :

Pierre combat les inégalités et nous défend des injustices.

– Modification d'un adjectif ou d'un nom par une négation : *(moins, peu, pas, aucunement,*

pas le moins du monde) chanceux, heureux, etc. ; *(Sans (aucune+la moindre)+point de+pas de) espoir, réussite, victoire,* etc.

– Ambiguïté sémantique (pouvant avoir une valeur positive ou négative en fonction des contextes) : complice (dans le crime *vs.* dans la vie).

— Nous avons également pu constater que la précision du système augmenterait s'il était capable de reconnaître :

  – des expressions/mots composés (qui très souvent n'ont pas la même valence que les éléments lexicaux dont ils sont constitués) : *centre dramatique, art dramatique, coup de foudre, etc.*

  – des noms propres à forte connotation : Front National, Brigades Rouges, etc.

  – des expressions figées ou métaphoriques comme *Au 70e de ses meetings, Jean fait un tabac à l'applaudimètre* ou *Pierre à la paupière qui papillote de joie, il est tendu d'impatience ...*

Par ailleurs, rappelons que nous souhaitons évaluer les contextes phrastiques dans le but d'établir si les noms identifiés sont évoqués dans un environnement plutôt positif ou plutôt négatif. Il est évident qu'en fonction de la place que ces noms occupent dans la structure syntaxique de la phrase, leur relation au contexte sera variable. Ceci est particulièrement remarquable dans le cas de phrases à incises : dans la phrase *« P », dit Jean*. Le nom *Jean* apparaît dans une incise que l'on peut considérer comme une parenthèse discursive. Le discours rapporté « P » et l'incise *dit Jean* appartiennent à des niveaux discursifs différents : l'incise est une intervention de l'auteur du texte qui attribue les propos rapportés à une personne. Il n'est donc pas nécessairement approprié d'appliquer les tests lexicaux sur « P » pour évaluer la situation du sujet de l'incise. En d'autre termes, une personne peut tenir des propos très marqués positivement ou négativement sans que ces propos ne la concernent directement. Il en va de même avec d'autres incises du type : *aux dires de Jean, selon Jean, d'après Jean, pour Jean, si l'on en croit Jean...*

Selon Jean, le trou du Lyonnais serait de l'ordre de 300 milliards de francs belges.

5. Conclusion

Nous avons constitué un corpus de référence permettant d'évaluer et de comparer des techniques de calcul automatique de la valence affective de phrases. Ce corpus nous a permis de mesurer la pertinence et les limites d'une approche basée sur le lexique. Il est mis librement à la disposition des chercheurs[5] qui souhaiteraient confronter leur système d'analyse aux résultats de l'évaluation humaine.

Les résultats obtenus (corrélation pouvant aller de 0.62 à 0.83 avec les jugements humains) sont relativement satisfaisants,particulièrement dans l'évaluation d'unités textuelles aussi réduites que des phrases.

A ce stade, une limite évidente de notre système de mesures provient du fait que nous n'effectuons aucune analyse linguistique des phrases pour prendre en compte des phénomènes syntaxiques (la présence d'une négation, d'incises, etc.) ou lexicaux (mot composés à valeur négative : extrême droite) qui altère la valeur affective des phrases. Sans ambitionner une véritable « compréhension du texte » (qu'il serait d'ailleurs bien vain de chercher dans le cadre strict d'une phrase), nous pensons que l'intégration de règles linguistiques et la prise en compte

[5] Http://cental.fltr.ucl.ac.be/publi/2004/valenceaffective.html

d'unités lexicales composées devraient permettre d'améliorer les évaluations automatiques. Ce sera l'une des prochaines étapes de cette recherche. Nous réaliserons également l'interfaçage de notre programme avec un système de veille de la presse sur Internet[6] (Fairon 1999) dans le but de suivre l'évolution des contextes « affectifs » dans lesquels apparaissent les noms de personnalités et de proposer un baromètre politique actualisé automatiquement de jour en jour. L'implémentation d'un tel système demandera d'analyser autant que possible toutes les occurrences des noms de personnalités dans la presse et donc également de résoudre au maximum les problèmes liés à la présence d'anaphores dans les articles.

Références

Anderson C.W. and McMaster G.E., (1982). Computer assisted modeling of affective tone in written documents », *Computers and the Humanities*, Vol. 16: 1-9.

Bestgen Y. (1994). Can emotional valence in stories be determined from words ?, *Cognition and Emotion*, Vol. 8: 21-36.

Bestgen, Y. (2002). Détermination de la valence affective de termes dans de grands corpus de textes. *Actes du Colloque International sur la Fouille de Texte 2* (pp. 81-94). Nancy : INRIA.

Das S. and Chen M., Yahoo! for Amazon : Opinion extraction from small talk on the web, Working Paper (under review), Décenbre 2001, Santa Clara University.

Fairon C. (1999). Parsing a Web site as a corpus, In C. Fairon (ed.). *Analyse lexicale et syntaxique: Le système INTEX*, Lingvisticae Investigationes Tome XXII (Volume spécial), Amsterdam/Philadelphia: John Benjamins Publishing, 450 p.

Fairon, C. and Watrin P. (2003). From extraction to indexation. Collecting new indexation keys by means of IE techniques, EACL 2003, Workshop on Finite State Methods. Budapest

Heise D.R. (1965). Semantic differential profiles for 1000 most frequent english words, *Psychological Monographs*, Vol. 79: 1-31.

Hogenraad R. and Bestgen Y. (1989). On the thread of discourse : Homogeneity, trends, and rhytms in texts, *Empirical Studies of the Arts*, Vol. 7: 1-22.

Hogenraad R., Bestgen Y. and Nysten J.L. (1995). Terrorist Rhetoric : Texture and Architecture, In Nissan et Schmidt (Eds.), *From Information to Knowledge*, Intellect, p. 48-59.

Pang, B., Lee, L. and Vaithyanathan, V. (2002). Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of the 2002 Conference on Empirical Methods in natural language processing* (pp. 79-86)

Polanyi, L. and Zaenen, A. (2003). Shifting attitudes. In L. Lagerwerf, W. Spooren, & L. Degand (Eds.), *Determination of information and tenor in texts: Multidisciplinary approaches to discourse 2003* (pp. 61-69). Münster: Nodus Publikationen.

Schmidt, H. (1994). Probabilistic Part-of-Speech Tagging Using Decision Trees. Version électronique disponible sur [http://www.ims.uni-stuttgart.de/Tools/DecisionTreeTagger.html].

Turney, P. and Littman, M. (2002). Unsupervised learning of semantic orientation from a hundred-billion-word corpus. Technical Report, National Research Council Canada.

Whissell C.M., Fournier M., Pelland R., Weir D. and Makarec K., (1986). A dictionary of affect in language : IV. Reliability, validity, and applications, *Perceptual and Motor Skills*, Vol. 62: 875-888.

Wilks Y. (1997). Information Extraction as a Core Language Technology, in M. T. Paziensa (Ed.) *Information Extraction*, Springer, p. 1-9.

[6] Http://glossa.fltr.ucl.ac.be