

Étude de la variabilité axiologique en fonction du contexte linguistique

soumission *CILPR XXVII* - Section 10 - Noémi Boubel

Notre recherche a pour cadre l'étude de l'expression d'une opinion ou d'une évaluation dans un énoncé. La question est, en particulier, de savoir si celui-ci est positif ou négatif. D'une façon naïve, on pourrait avancer que la polarité d'un texte est véhiculée par des items lexicaux positifs ou négatifs (*excellent, horrible, etc.*). La réalité linguistique est bien plus complexe ; la polarité globale d'un énoncé s'avère fortement dépendante du contexte. En effet, un contexte particulier peut modifier (annuler, inverser, atténuer, intensifier) la polarité initiale d'un terme. Ainsi, il n'est pas rare d'utiliser des termes positifs pour exprimer une évaluation négative et inversement, comme on peut le voir dans les exemples (1), (2), (3) et (4).

1. *Touchant de naïveté, hilarant de médiocrité, une vraie curiosité difficilement conseillable.*
2. *Ca aurait pu être brillant, c'est juste navrant (...).*
3. *Avec de tout petits moyens qui ne handicapent jamais vraiment son film, (...)*
4. *(...) un film aride, peu loquace, aussi chaleureux que l'Allemagne de l'Est de la grande époque.*

La question de recherche que nous abordons ici concerne précisément cet aspect : dans quels cas la valeur d'un item lexical positif ou négatif est inversée ou annulée par son contexte ? Pour répondre à cette question, nous proposons une méthodologie permettant d'analyser de façon systématique des productions réelles qui présentent cette situation (un item positif dans un énoncé négatif et inversement). Nous traitons cette problématique sous l'angle de la linguistique de l'énonciation et nous nous basons en particulier sur les théories linguistiques de l'expression de la subjectivité et de l'évaluation afin de définir une grille d'analyse pour l'étude de corpus. Nous détaillons ces deux aspects ci-dessous.

La méthodologie que nous proposons se base, d'une part, sur un corpus d'extraits de critiques de films provenant de presse traditionnelle ou en ligne. Ces critiques sont classées selon leur note (de très négatif à très positif) et seuls les passages particulièrement représentatifs d'une opinion ont été gardés¹. D'autre part, une liste d'items lexicaux positifs ou négatifs est établie. Nous nous proposons d'effectuer une analyse systématique d'un échantillon d'énoncés négatifs contenant un item positif, et d'énoncés positifs contenant un item négatif.

L'annotation systématique des énoncés retenus se base sur les concepts élaborés en linguistique de l'énonciation qui étudie la trace de l'inscription du sujet parlant dans l'énoncé. Kerbrat-Orecchioni (1997), en particulier, aborde cette question sous l'angle de l'expression de la subjectivité et s'attache à décrire notamment l'expression d'une *évaluation* (portée par des *subjectivèmes*) dans un énoncé. C'est à ce niveau de subjectivité qu'elle introduit la notion d'*axiologie*, qui regroupe les éléments « porteurs d'un trait évaluatif bon/mauvais » et qui peut donc se rapprocher de notre distinction positif/négatif.

La *variabilité axiologique* est largement traitée dans cette approche, dans la mesure où l'on considère qu'une étude de la langue hors situation d'énonciation n'a pas de sens et que la détection automatique d'une polarité n'est qu'hasardeuse (Kerbrat-Orecchioni, 1997, p.87). Le principe d'inversion, d'annulation, ou d'intensification d'une valeur axiologique n'est alors pas traité en tant que tel. Nous cherchons ici à mettre en évidence, dans ces théories linguistiques, certains traits évaluatifs qui ont une influence sur la valeur axiologique d'un subjectivème, sans être eux-mêmes axiologiques.

Kerbrat-Orecchioni (1997) postule l'existence de deux grands types d'évaluatifs : (1) des *axiologiques* (axe positif/négatif), mais aussi (2) des *modalisateurs* porteurs d'une valeur de vérité (vrai/faux). Ces modalisateurs vont être particulièrement importants dans notre perspective dans la mesure où le sens global (et notamment l'axiologie) d'un propos dépend de l'attitude de l'énonciateur par rapport à ce propos.

Charaudeau (1992) va plus loin et dresse une liste de modalités spécifiques à l'expression d'une évaluation. Il distingue notamment deux types de modalités : les *modalités d'évaluations logiques* (qui portent sur un axe vrai/faux) et les *modalités d'évaluations axiologiques*. Galatanu (2000) complète ce modèle en décrivant comment plusieurs modalités peuvent se combiner. Vernier (2011), dans une perspective de détection automatique, met en évidence que ces phénomènes peuvent influer sur l'axiologie d'autres subjectivèmes. Ainsi, les modalités d'évaluations logiques peuvent annuler voire inverser l'axiologie d'un item positif ou négatif (ex. 5). D'autre part, deux modalités (logiques ou axiologiques) peuvent influer l'une

¹ Ce corpus a été constitué sur la base du site allocine.fr qui restitue l'avis de différents journaux sur un film sous la forme de très brefs extraits des articles en question : <http://www.allocine.fr/film/fichefilm-135528/critiques/presse/>

sur l'autre lorsqu'elles se combinent (ex. 6)². Enfin, il est important de citer l'impact éventuel d'une structure syntaxique de négation sur des évaluatifs.

5. *Je ne pense pas que Domenech soit le sélectionneur légitime de cette équipe de France.*

6. *J'aime sa façon de mentir.*

Compte tenu de ces différents constats, nous choisissons de baser notre grille d'analyse sur les deux concepts clés qui viennent d'être décrits : (1) les traits axiologiques et (2) les modalisateurs et modalités logiques. Nous mettons ensuite en évidence comment ces différentes unités se combinent pour aboutir à la valeur axiologique globale de l'énoncé étudié. Nous prenons également un soin particulier à l'analyse des structures syntaxiques de négation : leur forme, leur impact éventuel sur un subjectivème, etc.

Notre présentation aura pour but d'exposer la méthodologie adoptée, les choix d'annotation ainsi que les problèmes rencontrés, avant de fournir les différents résultats qualitatifs et dans une certaine mesure quantitatifs issus de l'annotation. Les énoncés particuliers que sont les critiques journalistiques de films sont intéressants à plus d'un titre, car ils présentent un large panel de stratégies complexes pour exprimer une évaluation. D'une part, la trace de l'énonciateur y est rarement explicite, dans la mesure où l'on rencontre très peu d'embrayeurs de personnes. Cela permet d'accéder à des types d'évaluations (subjectivité de 2e niveau) plus délicates à détecter, portées par des phénomènes discursifs, lexicaux et syntaxiques variés. Ainsi, notre méthodologie permet de mettre en lumière certaines stratégies utilisées pour inverser ou annuler l'axiologie d'un subjectivème : combinaison de deux termes d'axiologie opposée (ex. 1), structures conditionnelles (ex. 2), structures grammaticales de négation (ex. 3), structures d'opposition-concession, etc. D'autre part, la variabilité axiologique d'un terme est fréquente dans ce type d'énoncé. Les évaluations se basent en effet souvent sur des connaissances du monde partagées ou des codifications sociales. Ainsi, dans l'exemple (4), l'inversion de l'axiologie positive de « *chaleureux* » passe par un processus d'ironie basé sur la connotation clairement négative de l'expression « *l'Allemagne de l'Est de la grande époque* ».

Pour conclure, notons que cette étude s'inscrit dans un but applicatif plus large : classer automatiquement des textes ou des phrases selon leur caractère positif ou négatif. Il s'agit d'un domaine de recherche à part entière du traitement automatique des langues communément appelé *fouille d'opinion* ou *analyse de sentiments* (Pang et Lee, 2008 ; Liu, 2012). La problématique du rôle du contexte dans ce domaine a donné naissance à la notion de *modifieur* (*contextual valence shifters* ou *modifiers*), notion qui désigne un phénomène contextuel (lexical, syntaxique, voire discursif) ayant un impact sur la polarité *a priori* d'un item lexical. Les travaux qui s'attachent à établir une liste de ces modificateurs (Polanyi et Zaenen, 2006 ; Taboada et al., 2011) cherchent à modéliser au mieux ces phénomènes dans une perspective d'automatisation, mais une analyse linguistique basée sur corpus ou sur des théories linguistiques n'est pas réellement menée. Sur cette constatation, notre approche cherche à ménager deux contraintes : (1) la contrainte de simplification inhérente à la modélisation d'un phénomène dans une perspective d'automatisation, (2) la nécessité de s'approcher au plus près des usages linguistiques réels et de rendre compte au mieux des stratégies d'expression d'une opinion ou d'une évaluation.

Bibliographie

- Charaudeau, P. (1992). *Grammaire du sens et de l'expression*. Hachette Education.
- Galatantu O. (2000). « Signification, sens, formation ». *Education et Formation, Biennales de l'éducation*, (sous la direction de Jean-Marie Barbier, d'Olga Galatantu), Paris : PUF.
- Kerbrat-Orecchioni, C. (1997). *L'énonciation : de la subjectivité dans le langage*. Colin.
- Liu, B. (2012). *Sentiment Analysis and Opinion Mining*. Morgan & Claypool Publishers.
- Pang, B. et Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), 1–135.
- Polanyi, L. et Zaenen, A. (2004). Contextual valence shifters. *Proceedings of AAAI Spring Symposium on Exploring Attitude and Affect in Text*, pages 106–111.
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K. et Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational Linguistics* 37(2) :267–307.
- Vernier, M. (2011). *Analyse à granularité fine de la subjectivité*. Thèse de doctorat en Informatique, Université de Nantes.

2 Exemples issus du corpus Blogoscopie, cités dans Vernier (2011)