

Avis sur les restaurants en ligne: détection et analyse quantitative

Hyun Jung KANG

Orange Innovation (Lannion)

The study examines, from various perspectives, restaurant evaluations written by customers and published online. From a manual observation of the corpus, we propose a model to study evaluations of experiences in restaurants; we show that evaluations can have several functions and contain different types of information: reviewers write reviews to give opinions, make suggestions, express intentions, and describe their experiences. We conduct various quantitative analyses using corpus linguistics methods which offer the possibility to identify the keywords of the reviews as well as their grammatical and topological characteristics, including keyword analysis, the distribution of parts of speech and the positions of evaluation categories within reviews. We finally try to automatically detect the evaluation categories in our corpus. We test and validate a supervised learning approach based on our model completed with other linguistic and textual features.

1. Introduction

Depuis l'avènement du web 2.0, la problématique de l'analyse des évaluations publiées en ligne suscite un grand intérêt dans les domaines de la linguistique et du traitement automatique des langues (TAL). En linguistique française, l'évaluation est étudiée à différents niveaux: (i) au niveau morphologique (Amiot 2012; Amiot & Stosic 2015; Fradin 2003; Lefer & Grabar 2013, 2014), (ii) au niveau lexical (Cheminée 2006; Mathieu 2006; Pupier 1998; Vernier & Monceaux 2010), (iii) au niveau des séquences lexico-grammaticales (Legallois & Ferrari 2006), (iv) au niveau des séquences discursives (Benamara et al. 2016) et (v) au niveau énonciatif (Charaudeau 1992; Gosselin 2010, 2017). La linguistique anglophone fait également appel à un large éventail de notions: théorie de l'*appraisal* (Martin & White 2003), *stance* (Biber & Finegan 1989), évaluation (Hunston & Thompson 2000) et *nonveridicality* (Taboada & Trnavac 2013).

En TAL, c'est à travers la tâche de fouille d'opinions que ce domaine s'intéresse à l'évaluation. Son objectif est de développer un système qui permet de reconnaître et d'extraire les expressions évaluatives ainsi que leur polarité (Liu 2015). Les recherches en TAL portant sur la fouille d'opinions prennent comme point de départ la constitution de ressources lexicales. Hatzivassiloglou et McKeown (1997), par exemple, associent les adjectifs et les groupes adjectivaux à une polarité, ce qui permet ensuite de donner un score au document. Les travaux évoluent progressivement vers une classification automatique des évaluations et réduisent la problématique de l'évaluation à une question de classification de polarité (positive/négative/neutre) ou de subjectivité (subjective/objective). Pour ce faire, les techniques d'apprentissage supervisé, d'apprentissage non supervisé et plus récemment d'apprentissage profond sont exploitées.

Ces dernières années, les recherches sur la fouille d'opinions se sont réorientées vers l'extraction de suggestions. À titre d'exemple, l'une des tâches de l'atelier SemEval-2019¹ (Negi et al. 2019) concernait l'extraction de suggestions dans des avis et forums en ligne. La notion de suggestion est abordée sous différentes formes. Le travail de Goldberg et al. (2009), qui utilise le terme de "souhait" (*wish*), apparaît fondateur dans le domaine de la détection automatique de suggestions. De leur côté, Ramanand et al. (2010) proposent de distinguer deux types de souhaits: celui visant à améliorer le produit (ce que Brun et Hagège (2013) nomment "suggestion d'amélioration") et celui visant à l'acheter. Wicaksono et Myaeng (2012, 2013) emploient le terme de "conseil" (*advice*) dans le contexte de conseils de voyages publiés sur des forums de discussion. Negi et Buitelaar (2015) se penchent sur les suggestions de client à client (*customer-to-customer* ou CTC), où un client cherche à aider les autres, ainsi que sur les suggestions d'amélioration du produit ou du service adressées aux fournisseurs. Ainsi, l'intérêt que les chercheurs portent aujourd'hui à la fouille d'opinions dépasse largement la notion d'opinion positive ou négative.

Cette étude porte sur la modélisation de l'évaluation dans le domaine des avis rédigés en ligne, en français, par des clients de restaurants. Le texte articule deux approches à la fois empiriques et méthodologiques. Nous nous inscrivons dans une démarche de linguistique de corpus et de TAL: en partant d'une description des procédures mises en place dans l'appareillage expérimental élaboré, nous cherchons à rendre explicites les propriétés du corpus. L'observation manuelle du corpus constitué montre que les évaluations peuvent avoir plusieurs fonctions et contenir différentes informations. Une modélisation de l'évaluation est ensuite proposée sous forme de typologie comprenant six catégories: opinion positive, opinion négative, opinion mixte, intention, suggestion et description. Le corpus est annoté sur la base de ces catégories, ce qui permet de l'analyser d'un point de vue quantitatif en exploitant les méthodes de la linguistique de corpus anglo-saxonne. En outre, une démarche lexicométrique fondée sur les méthodes d'analyse des données de masse est intégrée à l'analyse quantitative. Dans une perspective de TAL, un module de détection automatique est développé pour l'annotation automatique des évaluations.

2. Corpus étudiés

Les données traitées sont composées d'évaluations collectées sur un site Internet² proposant des fiches détaillées de restaurants ainsi que des avis de clients. Seuls les clients qui ont effectué une réservation sur le site et qui ont consommé dans le restaurant donné peuvent laisser un avis. De plus, au

¹ SemEval 2019 tâche 9, <https://competitions.codalab.org/competitions/19955> [consulté le 25 octobre 2021].

² <https://www.lafourchette.com> [consulté le 25 octobre 2021].

moment de la collecte du corpus, le site disposait d'une plus grande quantité d'avis récents que les autres sites. Pour cette raison, ce site a été choisi pour l'étude. Les avis sont exprimés en texte libre et à travers une notation. Afin d'assurer l'homogénéité du corpus, nous avons sélectionné des restaurants situés à Paris et proposant une cuisine traditionnelle française. Les avis sur les restaurants ont également été choisis en fonction de leur date de parution: pour chaque restaurant, les avis antérieurs au mois de novembre 2016 ont été extraits, dans la limite de 200 avis³. Nous avons choisi 87 restaurants, correspondant à 6 287 avis et 17 268 phrases, avec une moyenne de dix mots par phrase.

3. Modéliser les évaluations

Le travail présenté vise à étudier sous différents angles des évaluations de restaurants publiées par des clients, en faisant appel à des méthodes de la linguistique de corpus et du TAL. Pour atteindre cet objectif, un modèle conceptuel a été développé, qui décrit les différents types d'évaluation observés dans le corpus. Ce modèle s'appuie sur quatre catégories: l'opinion (positive/négative/mixte), la suggestion, l'intention et la description. Dans le but d'étudier et de traiter automatiquement l'évaluation, nous avons procédé à l'annotation manuelle des phrases selon le modèle élaboré.

3.1 Typologie

3.1.1 Opinion

La notion d'"opinion" s'articule autour de deux dimensions, épistémique et axiologique (Zhang 2012). Les réflexions menées dans le cadre de cette étude nous ont amenée à nous limiter à la dimension axiologique de l'opinion. Dans les données d'avis sur les restaurants, les clients mobilisent des valeurs en exprimant ce qui est, selon eux, bon ou mauvais, agréable ou désagréable, satisfaisant ou insatisfaisant, souhaitable ou à éviter. Dans cette étude, les opinions représentent donc les idées que le consommateur se fait du restaurant et la façon dont il qualifie ce dernier. Si l'interprétation axiologique de la valeur accordée à l'entité donnée se réfère à une qualification favorable, il s'agit d'une opinion positive; si l'interprétation relève d'une appréciation défavorable, l'opinion est négative. Lorsque l'opinion comporte les deux polarités (positive et négative), elle est qualifiée d'opinion mixte.

Différents marqueurs ayant un impact sur la polarité ont été observés dans le corpus traité. Le lexique évaluatif rend explicite la polarité de l'opinion. À titre d'exemple, *bon*, *jolie* et *chaleureux* sont des marqueurs de l'opinion positive (1), alors que *médiocre*, *bruyant* et *cher* sont plutôt des marqueurs de l'opinion

³ Le nombre d'avis a été déterminé en tenant compte de la quantité disponible d'avis exprimés pour chaque restaurant.

négative (2). Par ailleurs, lorsque la négation (*ne...pas*, *aucun* et *peu*) est appliquée aux mots ayant une valeur positive, le mot positif est placé dans la zone négative, et inversement (3). La polarité des opinions est également déterminée grâce au lexique des émotions ou des sentiments, ces derniers étant révélés au cours de l'expérience au restaurant (2, 4). Par ailleurs, certains clients manifestent leur satisfaction ou leur mécontentement à travers une représentation factuelle de la situation concernée, jugée favorable ou défavorable selon les stéréotypes présents dans la société (5).

- 1) *Plat très **bon**, présentation très **jolie**, accueil **chaleureux**!*
- 2) *Nous avons été **déçus** par une sole de **médiocre** fraîcheur.*
- 3) Le service **n'est pas** à la hauteur!
- 4) *Nous sommes rentrés **ravi** de notre soirée.*
- 5) *Réserve confirmée à 21h 30, 45 minutes d'attente.*

Cependant, la notion de polarité est complexe. Plus précisément, la polarité de certains mots tels que *courte*, *simple*, *petit*, *classique* ou *surprenant* est ambiguë. Une carte *courte* ou *simple*, par exemple, peut être rassurante selon les lecteurs: le recours à peu d'ingrédients implique que ceux-ci sont probablement de meilleure qualité. Toutefois, dans le même temps, cette carte limite la variété des choix. De plus, la polarité d'une phrase peut être modifiée en fonction des modificateurs (ou *valence shifters*, Polanyi & Zaenen 2006) tels que les intensifieurs, les atténuateurs, la négation ou les connecteurs. Dans cette étude, le connecteur *mais* joue un rôle important dans la détermination de l'opinion mixte (6): il hiérarchise l'importance à accorder aux deux propositions et peut donc modifier la valeur de la conclusion à tirer de la phrase.

- 6) *Plat **excellamment** cuisiné, **mais** certaines entrées **décevantes**.*

Dans les écrits numériques, la polarité et l'intensité se manifestent également au moyen d'autres éléments (Vasquez 2014), comme les émoticônes (7), les ponctuations multiples en cascade (8) ou les mots en majuscules (9).

- 7) *Nous avons passé une merveilleuse soirée :)*
- 8) *Vraiment exceptionnel!!!*
- 9) *Ce fut vraiment un **REGAL**!*

3.1.2 Suggestion

Les différents objectifs visés par le client sont mis en évidence par la double nature des destinataires des avis en ligne. Ainsi, dans notre étude, la suggestion vise, d'une part, une amélioration relative des produits (ou services) et, d'autre part, s'adresse aux lecteurs dans le but de leur donner des conseils. Dans cette optique, les suggestions sont destinées à la fois aux restaurateurs (afin de leur

faire prendre conscience des problèmes) et aux autres clients potentiels (futurs clients).

La suggestion proposée par les clients se manifeste par divers marqueurs: les verbes de parole désignant l'ordre ou le conseil (10), le mode conditionnel (11), le mode impératif (12, 13), les pronoms personnels ou possessifs et les adjectifs possessifs de la deuxième personne (13).

- 10) *Je **recommande** vivement.*
- 11) Un petit niveau musical en fond sonore **aurait été** un plus.
- 12) ***Allez** dans ce restaurant les yeux fermés!*
- 13) *Si **vous** aimez les fruits de mer, **n'hésitez pas** une seconde.*

3.1.3 Intention

Dans cette étude⁴, l'intention renvoie au fait que les clients communiquent leur souhait de retourner ou non dans un restaurant. Cette intention peut être considérée comme positive ou négative. L'intention positive implique souvent le souhait de revenir au restaurant; il s'agit donc d'une évaluation positive. En revanche, l'intention négative manifeste le désir de ne pas répéter l'expérience; c'est donc une évaluation négative. L'intention représente ainsi un engagement volontaire du client, initié par lui.

L'intention est marquée explicitement dans le corpus à travers des verbes au futur ou comprenant le préfixe *re-* qui indique généralement la réitération d'une action (*revenir, retourner, refaire et renouveler*). Notons également que dans les avis observés, l'intention se retrouve majoritairement dans la dernière phrase. Celle-ci résume ainsi l'ensemble de l'expérience gastronomique en donnant une évaluation globale du restaurant, comme dans les exemples suivants:

- 14) *Un lieu très agréable et nous y **retournerons!***
- 15) *Je **n'y reviendrai pas.***

3.1.4 Description

Les trois catégories précédentes (l'opinion, la suggestion et l'intention) représentent l'évaluation exprimée par un client. La description, en revanche, renvoie aux informations factuelles associées à l'expérience vécue. Ce type d'information est peu reconnu dans la fouille d'opinions; il a même été considéré comme inutile dans l'étude menée par Ganu et al. (2009). La description est liée, d'une part, à des informations factuelles ayant un impact avant la visite du

⁴ La notion d'intention a été considérablement abordée dans le domaine de l'intelligence artificielle. Le modèle logiciel croyance-désir-intention (Belief-Desire-Intention ou BDI) modélise le comportement d'un agent intelligent, inspiré de la théorie de Bratman (1987). Selon ce modèle, l'intention désigne l'action que l'agent entreprend dans le but de satisfaire son désir.

restaurant, telles que le motif de cette dernière (16) ou des précisions concernant une réservation effectuée (17). D'autre part, elle est relative à des informations factuelles concernant l'expérience vécue lors de la visite, comme le(s) plat(s) choisi(s) ou la composition du plat (18). La description fait donc intervenir les tendances, les modes de vie et le quotidien des clients, et exerce des influences sur la perception, l'attitude et le comportement des lecteurs. Ainsi, ce type d'information est étroitement lié au contexte de l'expérience vécue, permettant de mieux cerner chacun des clients.

16) *J'y suis allé pour faire ma demande en mariage.*

17) *Nous avons réservé lundi soir, une table de 4.*

18) *Nous avons pris en entrée le foie gras poché.*

3.2 Annotation manuelle

Afin d'explorer l'évaluation d'un point de vue quantitatif et de nourrir un système d'apprentissage automatique, nous avons procédé à l'annotation manuelle des avis selon la typologie élaborée et présentée précédemment. La tâche d'annotation consistait à attribuer à chaque phrase une étiquette parmi des catégories prédéfinies: POS_OPINION (opinion positive), NEG_OPINION (opinion négative), MIX_OPINION (opinion mixte), SUGGESTION, INTENTION et DESCRIPTION.

Pour ce faire, le corpus a été segmenté en phrases en fonction des signes typographiques (. ! ?) puis annoté selon la typologie d'évaluation. Le nombre d'avis à annoter s'élevait à 6 287, pour un total de 17 268 phrases. Chaque phrase a donc été annotée manuellement comme illustré dans le tableau 1.

Phrase	Catégorie
<i>Nous nous sommes tous régalez.</i>	POS_OPINION
<i>Les desserts m'ont déçu.</i>	NEG_OPINION
<i>L'addition est corsée mais tout est top.</i>	MIX_OPINION
<i>Un léger fond musical serait un plus.</i>	SUGGESTION
<i>On reviendra!</i>	INTENTION
<i>Nous y étions pour les 30 ans de ma conjointe.</i>	DESCRIPTION

Tableau 1: L'exemple d'annotation manuelle

Dans les cas d'ambiguïté où plusieurs catégories étaient pertinentes, les catégories les moins représentées dans le corpus, comme l'intention ou la suggestion, ont été privilégiées. Nous justifions ce choix par des études en psycholinguistique (Carreiras et al. 1995; Gernsbacher 1990; Kim et al. 2004) qui affirment que le contenu présenté au début de la phrase semble avoir un effet plus marquant dans la détermination du jugement final. Ainsi, la phrase (14)

qui peut être assignée à deux catégories (opinion positive et intention), a été annotée comme intention. La phrase "*Adresse à recommander, nous y retournerons.*", par exemple, a ainsi été annotée comme suggestion.

Avant de procéder à l'annotation manuelle finale, nous avons validé la typologie proposée au moyen de l'accord inter-annotateur (AIA), qui permet d'assurer un bon niveau d'annotation. Dans cette optique, un test d'annotation a été réalisé sur un extrait du corpus (100 phrases choisies de manière aléatoire mais en respectant la distribution de chaque catégorie) par trois doctorants en linguistique. L'AIA entre ces annotateurs a été calculé en appliquant la mesure Kappa de Fleiss (Fleiss 1981). L'accord porte sur l'attribution d'une même étiquette par chaque annotateur. Le score obtenu⁵ (0,90) est considéré comme "presque parfait" d'après l'échelle de Landis et Koch (1977). Il démontre que nos catégories et nos règles d'annotations sont bien définies et fiables. La typologie élaborée ayant été jugée pertinente, le reste du corpus (17 168 phrases) a été annoté par l'un des annotateurs, avec pour résultat un total de 17 268 phrases.

La répartition des catégories (figure 1) s'avère fortement déséquilibrée: l'opinion positive constitue la majorité des données (68,05 %), alors que les intentions et les descriptions n'en représentent que 4,86 % et 1,85 %, respectivement. Ce déséquilibre a été pris en compte lors de la détection automatique (voir section 5).

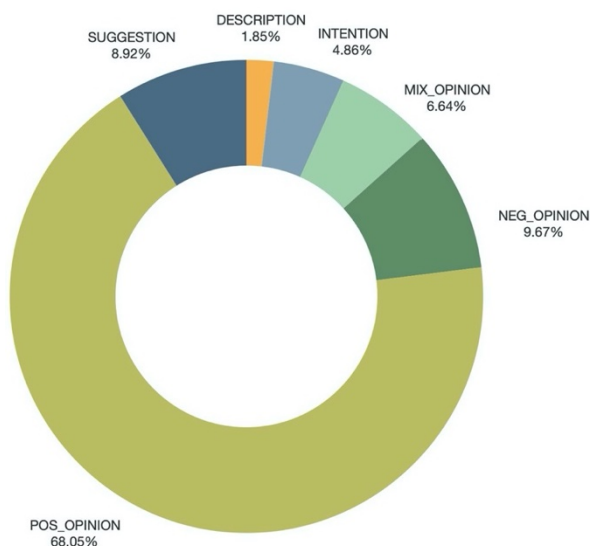


Figure 1: Répartition des catégories d'évaluation

4. Analyse quantitative

Le modèle d'évaluation proposé est justifié non seulement par les observations directes sur corpus, mais aussi par sa projection sur les données. Le corpus annoté a permis d'effectuer différentes analyses quantitatives à l'aide des

⁵ Nous avons suivi le calcul illustré dans Pustejovsky & Stubbs (2012).

méthodes de la linguistique de corpus. Pour chaque catégorie d'évaluation, une analyse quantitative a d'abord été réalisée sur les mots-clés extraits. Nous avons ensuite procédé à l'étude de la distribution des parties du discours, suivie de l'examen de leur localisation au sein d'un avis. Les analyses ont été effectuées à l'aide d'outils tels que le langage Python, le logiciel R⁶ et le logiciel AntConc⁷ (Anthony 2019).

4.1 Nettoyage et prétraitement

Avant de réaliser les analyses quantitatives et automatiques, les données ont été nettoyées et prétraitées afin de réduire la variation des mots à indexer lors du traitement automatique. Ce processus est essentiel pour explorer des données provenant de communications médiées par l'ordinateur (Poudat et al. 2020), car ces dernières possèdent des particularités qui les distinguent de celles de l'écrit standard: l'utilisation de majuscules pour donner de l'emphase (*EXCELLENT*), d'émoticônes (":)", ":(") et d'abréviations (*resto* au lieu de *restaurant*) ou encore les fautes d'orthographe ou de frappe (*acceuil* au lieu de *accueil*).

Les traitements effectués dans cette étude sont les suivants: le passage des mots en minuscules, le remplacement des chiffres par "NUM", des symboles € par "euros" et % par "pourcent", des émoticônes par "emoPOS" ou "emoNEG" selon la polarité⁸, ainsi que la normalisation des abréviations (par exemple, *resto* en *restaurant*, *déco* en *décoration*). Le corpus a ensuite été lemmatisé et étiqueté. Il s'agit d'assigner à chaque mot du corpus son lemme, c'est-à-dire sa forme canonique, et d'ajouter des annotations supplémentaires sur sa catégorie morphosyntaxique. Il existe différents étiqueteurs pour réaliser ces traitements, dont spaCy⁹, Treetaggerwrapper¹⁰ et StanfordCoreNLP¹¹. D'une manière générale, les étiqueteurs sont très performants lorsqu'il s'agit de traiter l'anglais, mais moins pour le français; c'est notamment le cas de spaCy. En revanche, StanfordCoreNLP et Treetaggerwrapper produisent des résultats comparables. La différence porte sur une catégorie morphosyntaxique AUX (auxiliaire), absente chez Treetaggerwrapper. Pour cette raison, StanfordCoreNLP a été choisi pour ce travail. L'identification des auxiliaires permet de les éliminer dans les traitements et de ne se concentrer que sur les verbes plus significatifs. Par ailleurs, certains mots-outils ont été retenus dans cette étude car ils constituent des marqueurs ayant une valeur informative. À titre d'exemple, les embrayeurs

⁶ <https://www.r-project.org/> [consulté le 22 avril 2022].

⁷ <https://www.laurenceanthony.net/software/antconcl/> [consulté le 25 octobre 2021].

⁸ Les émoticônes peuvent se distinguer par la polarité: ceux qui ont une polarité positive comme ":", ":-)" et ceux qui sont plutôt de polarité négative ":(\" (Bifet et al. 2011). Une liste comprenant les émoticônes positifs et négatifs a été créée préalablement.

⁹ <https://spacy.io/models/fr/> [consulté le 22 avril 2022].

¹⁰ <https://treetaggerwrapper.readthedocs.io/en/latest/> [consulté le 22 avril 2022].

¹¹ <https://stanfordnlp.github.io/CoreNLP/download.html> [consulté le 25 octobre 2021].

de personnes – les pronoms personnels, pronoms possessifs et adjectifs possessifs de la première personne (*je, nous, moi, mon*, etc.) et de la deuxième personne (*vous, votre*, etc.) – renvoient au locuteur ou au(x) destinataire(s). La conjonction *mais*, quant à elle, joue un rôle important dans la détermination de l'opinion mixte, opposant les opinions positives et négatives au sein d'une même phrase. De plus, les adverbes qui expriment l'intensité, comme *très*, permettent de renforcer l'expression de l'opinion. Les chiffres, pour leur part, sont souvent employés dans la description pour indiquer des informations factuelles (le prix, le nombre de personnes, etc.). De ce fait, dans le cadre de l'analyse quantitative, nous retenons les lemmes des classes ouvertes (noms, verbes, adjectifs et adverbes), les embrayeurs de personnes, la conjonction *mais*¹² et les chiffres (NUM).

4.2 Analyse des mots-clés

La première analyse effectuée concerne les mots-clés (*keyword analysis*; Baker 2010), une pratique employée dans de nombreuses études linguistiques basées sur des corpus (McEnery 2016; Poole 2016; Scott & Tribble 2006; Seale et al. 2006). Elle consiste à comparer la fréquence relative des mots dans un corpus d'étude à celle observée dans un corpus de référence. Autrement dit, elle permet d'extraire les mots qui sont particulièrement fréquents dans le corpus d'étude et de révéler les caractéristiques saillantes et fonctionnelles de ce dernier. L'extraction de ces mots-clés a été effectuée à l'aide du logiciel AntConc (Anthony 2019) avec les paramètres par défaut¹³. Le logiciel s'appuie sur la mesure statistique du rapport de vraisemblance¹⁴ (*log-likelihood ratio*; Dunning 1993) afin de déterminer le "degré de saillance" (*keyness*). Les résultats de l'extraction des mots-clés sont présentés dans le tableau 2. Il s'agit des dix lemmes les plus saillants pour chaque catégorie.

¹² L'analyse de notre corpus montre que 57,5 % des opinions mixtes contiennent la conjonction *mais*. Cette dernière est d'ailleurs la conjonction la plus fréquemment observée dans le corpus. Nous avons donc fait le choix de ne retenir qu'elle.

¹³ Nous avons suivi les recommandations données dans la documentation de l'outil. <https://www.laurenceanthony.net/software/antconc/releases/AntConc343/help.pdf> [consulté le 25 octobre 2021].

¹⁴ Le rapport de vraisemblance est une mesure d'association qui consiste à quantifier la force d'une combinaison de mots (Charest et al. 2007).

POS_OPINION		NEG_OPINION		MIX_OPINION	
lemme	saillance	lemme	saillance	lemme	saillance
<i>très</i>	699.2	<i>peu</i>	356.52	<i>mais</i>	1319.01
<i>excellent</i>	442.77	<i>bémol</i>	305.46	<i>peu</i>	510.47
<i>accueil</i>	372.53	<i>pas</i>	294.84	<i>cher</i>	288.81
<i>bon</i>	365.49	<i>dommage</i>	286.32	<i>trop</i>	205.57
<i>service</i>	322.9	<i>ne</i>	255.84	<i>pas</i>	86.62
<i>parfait</i>	322.08	<i>seul</i>	243.31	<i>assez</i>	61.58
<i>agréable</i>	301.61	<i>trop</i>	234.91	<i>même</i>	58.34
<i>cuisine</i>	251.82	<i>NUM</i>	182.32	<i>élevé</i>	57.32
<i>cadre</i>	221.47	<i>cher</i>	161.5	<i>décevant</i>	47.87
<i>beau</i>	214.78	<i>déception</i>	108.97	<i>bruyant</i>	46.51

SUGGESTION		INTENTION		DESCRIPTION	
lemme	saillance	lemme	saillance	lemme	saillance
<i>recommander</i>	3075.7	<i>revenir</i>	1792.85	<i>NUM</i>	580.72
<i>vivement</i>	571.85	<i>retourner</i>	1459.54	<i>euro</i>	195.3
<i>conseiller</i>	356.33	<i>refaire</i>	565.15	<i>menu</i>	131.25
<i>vous</i>	308.22	<i>nous</i>	472.31	<i>nous</i>	102.37
<i>hésiter</i>	177.25	<i>je</i>	299.59	<i>prendre</i>	89.65
<i>aller</i>	129.99	<i>plaisir</i>	241.08	<i>anniversaire</i>	49.44
<i>découvrir</i>	111.06	<i>retenir</i>	166.74	<i>choisir</i>	47.2
<i>falloir</i>	106.12	<i>on</i>	148.34	<i>opter</i>	35.09
<i>y</i>	98.79	<i>bientôt</i>	132.67	<i>hasard</i>	30.42
<i>fortement</i>	93.65	<i>hâte</i>	111	<i>dessert</i>	27.19

Tableau 2: Mots-clés de chaque catégorie

Les mots saillants extraits montrent que les hypothèses sur les marqueurs linguistiques de chaque catégorie se vérifient. Les opinions positives sont associées au lexique évaluatif qui impliquent la polarité positive (*bon, parfait, agréable, beau*) et les cibles sur lesquelles elles portent (*accueil, service, cuisine, cadre*). Les opinions négatives se caractérisent par le lexique évaluatif négatif (*bémol, dommage, cher, déception*) et la négation (*peu, pas, ne*). Les opinions mixtes comprennent les lexiques relatifs aux opinions positives et négatives, et le terme le plus caractéristique de cette catégorie est la conjonction *mais*. Les suggestions sont liées aux verbes d'ordre et de conseil (*recommander* et *conseiller*). Elles se manifestent également à travers le pronom *vous*, qui construit un lien entre l'auteur et les lecteurs. Les intentions contiennent des verbes itératifs (*revenir, retourner, refaire*) et des pronoms (*nous, je, on*). La dernière catégorie, les descriptions, inclut des mots concernant les menus choisis (*menu, dessert, opter*), leur prix (*NUM, euro*) et

le motif de la visite (*anniversaire, hasard*). Nous constatons également que les noms et adjectifs sont majoritairement présents dans les opinions et les verbes dans les suggestions, intentions et descriptions. Les catégories d'évaluation sont caractérisées par des choix lexicaux/morphosyntaxiques que nous examinerons plus en détail dans la partie suivante, à l'aide d'une analyse factorielle des correspondances.

4.3 Distribution des parties du discours

Dans cette partie, nous examinons les caractéristiques morphosyntaxiques de chaque catégorie d'évaluation. Nous nous intéressons donc aux catégories morphosyntaxiques (ADJ, ADV, DET, NOUN, PRON et VERB). Dans cet objectif, une analyse factorielle des correspondances a été effectuée grâce au progiciel FactoMineR (Lê et al. 2008), disponible dans le langage R. Il est difficile de dégager manuellement les caractéristiques pertinentes de corpus volumineux; pour cette raison, certains chercheurs se tournent vers l'analyse factorielle des correspondances (AFC) (Desagulier 2020; Poudat & Landragin 2017). Présenté par Benzécri (1973, 1984), il s'agit d'une méthode qui s'applique aux tableaux de contingence, autrement dit, aux tableaux de comptage obtenus par le croisement de deux variables nominales. L'AFC permet d'étudier et de visualiser les relations existant entre les deux variables. Elle permet ainsi d'interpréter les rapprochements ou les distances entre les points d'un même ensemble par leur association avec ceux de l'autre ensemble (Lebart et al. 2019).

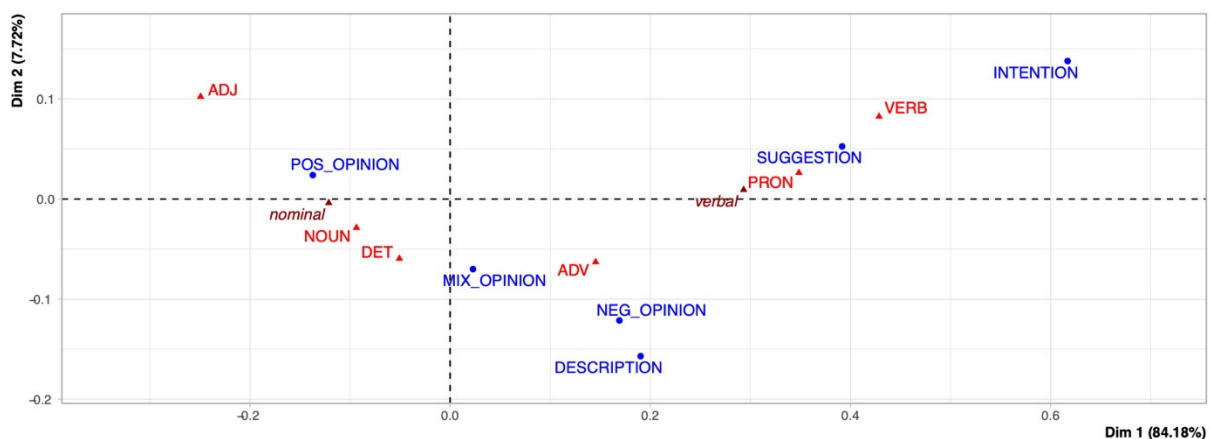


Figure 2: Analyse factorielle des correspondances

La figure 2 montre le lien entre les catégories morphosyntaxiques obtenues avec l'étiqueteur StanfordCoreNLP et les différents types d'évaluation. Les dimensions 1 et 2 expliquent respectivement 84,18 % et 7,7 % de l'inertie totale¹⁵, c'est-à-dire que les deux premières dimensions rendent compte à elles

¹⁵ La somme des variances des variables considérées (Baccini 2010).

seules de 91,9 % des variations observées dans nos données. Les catégories d'évaluation sont représentées par des points bleus et les catégories morphosyntaxiques par des triangles rouges. Les éléments supplémentaires ou illustratifs (Gower 1968) sont également projetés, qui regroupent des informations pertinentes: les valeurs totales de l'adjectif, du déterminant et du nom sont indiquées sous le nom "nominal" et celles de l'adverbe, du pronom et du verbe comme "verbal". Cependant, ces éléments ne contribuent pas à la constitution des dimensions, ils n'interviennent qu'a posteriori. Le centre du graphique correspond au profil moyen des deux variables. Un point donné s'écarte du centre de gravité s'il contribue aux axes davantage que le profil moyen. Les points en corrélation négative sont éloignés, tandis que ceux qui ont des profils similaires se situent de façon relativement proche.

En s'appuyant sur le premier axe, une opposition est observée entre l'opinion positive, qui se trouve à gauche du graphique, et les catégories restantes à droite. À proximité de l'opinion positive, nous observons les adjectifs, les noms et les déterminants, ce qui montre que les opinions positives tendent à préférer les catégories nominales: "Magnifique endroit, très bonne cuisine et accueil irréprochable!", "Très bon rapport qualité prix et service très aimable et efficace", "Cuisine originale avec de bons produits frais". En revanche, la suggestion fait intervenir pronoms et verbes, et l'intention recourt principalement aux verbes. La suggestion et l'intention sont toutes deux en lien avec l'action, comme nous l'avons constaté lors de l'observation manuelle et de l'analyse des mots-clés. La suggestion faite par un client implique une action future de la part du restaurateur et de ses futurs clients, et l'intention concerne l'action qui sera effectuée par le client lui-même. La description et l'opinion négative figurent en bas du graphique, illustrant que ces catégories sont davantage sensibles à l'opposition entre adjectifs et adverbes. Par ailleurs, leur position excentrée peut également s'expliquer par le fait que les opinions négatives et les descriptions sont formées d'un nombre plus restreint de segments, par rapport aux autres catégories (intention, suggestion, etc.). Enfin, l'opinion mixte, qui se situe à peu près au milieu du graphique, présente des caractéristiques de l'opinion positive et de l'opinion négative. Notons que l'AFC a permis de confirmer nos observations issues de l'approche par mots-clés.

4.4 Localisation des catégories d'évaluation

L'objectif de cette partie est de vérifier s'il existe un lien entre la catégorie d'évaluation et sa localisation dans un avis. D'après une observation manuelle, la description se manifeste souvent au début d'un avis, alors que la suggestion et l'intention ont tendance à apparaître vers la fin. Afin de confirmer cette observation, deux variables ont été envisagées: la catégorie d'évaluation et la position. Pour cette dernière, l'ouverture et la clôture (première et dernière phrases) ont été considérées comme les positions les plus marquantes. Par ailleurs, un graphique d'association (*association plot*) a été produit à l'aide du

logiciel R (figure 3). Ce graphique s'appuie sur les résidus normalisés de Pearson (Haberman 1973), exploités à l'aide du test du khi-deux d'indépendance. Ils permettent d'identifier les paires (catégorie d'évaluation et position) dont la fréquence est différente de la fréquence attendue si les catégories étaient distribuées de manière homogène au sein d'un avis. Autrement dit, les résidus de Pearson permettent d'identifier les associations privilégiées entre une catégorie d'évaluation et une position. Les avis composés d'une seule phrase n'ont pas été pris en compte. Le corpus étudié pour cette analyse comprend un total de 4 552 avis correspondant à 15 533 phrases.

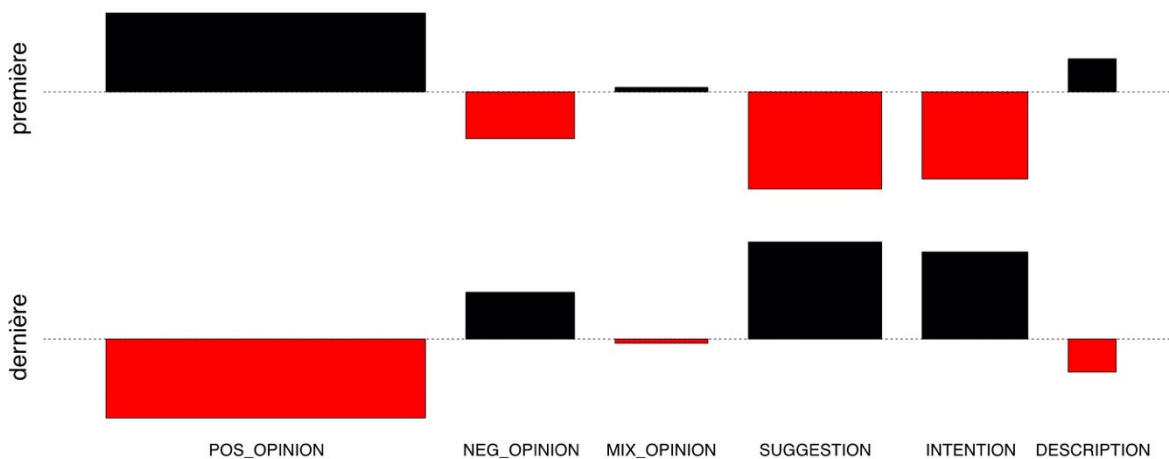


Figure 3: Graphique d'association entre la catégorie d'évaluation et la position

Le principe de lecture du graphique est le suivant: chaque cellule est représentée par un rectangle dont la hauteur est proportionnelle au résidu de Pearson, tandis que la largeur est proportionnelle à la racine carrée de l'effectif attendu. La surface de la cellule est donc proportionnelle à la différence entre la fréquence observée et la fréquence attendue. La ligne pointillée indique la ligne de référence: ainsi, si la valeur du résidu est positive (négative), la cellule se situe au-dessus (au-dessous) de cette ligne; si la fréquence observée d'une cellule est supérieure (inférieure) à sa fréquence attendue, la cellule apparaît au-dessus (au-dessous) de la ligne de référence, en noir (rouge).

Nous remarquons, dans la figure 3, que l'opinion positive et la description sont placées de manière statistiquement significative en première position d'un avis, tandis que la suggestion, l'intention et l'opinion négative sont placées en dernière position. Autrement dit, les clients commencent fréquemment leur avis par un arrière-plan (la description) ou directement par une opinion positive. En revanche, les avis ont tendance à finir sur des suggestions aux lecteurs (les clients potentiels et le restaurateur) ou l'expression d'un souhait de retourner ou non au restaurant (l'intention).

5. Détection automatique

La détection automatique de différentes catégories d'évaluation est fondée sur des méthodes couramment employées en TAL, telles que l'apprentissage supervisé. L'analyse du corpus réalisée dans la section précédente a permis de représenter les données le mieux possible et de prendre en compte leurs spécificités lors de la détection automatique. L'apprentissage supervisé est réalisé à partir du modèle entraîné sur les annotations manuelles de 17 268 phrases. Différentes expériences ont été réalisées dans le but de comparer les diverses méthodes de détection, décrites dans Eshkol-Taravella & Kang (2019) et Kang & Eshkol-Taravella (2020). Au cours des expériences, nous avons obtenu la meilleure moyenne pondérée de la F-mesure, la métrique d'évaluation courante en TAL, grâce au classifieur *support vector machine* (SVM) linéaire, donnant un score de 0,88. Nous nous contentons de présenter ici la synthèse de ces résultats.

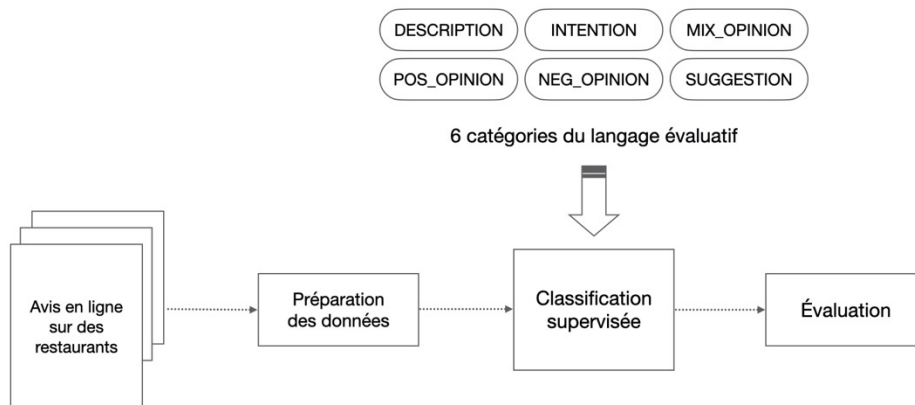


Figure 4: Schéma de la chaîne de traitements

Afin de détecter automatiquement l'évaluation modélisée dans le corpus d'avis, nous avons suivi plusieurs étapes (figure 4). Les avis ont d'abord été prétraités, segmentés et convertis en vecteurs. Le prétraitement des données textuelles consiste à proposer pour chaque mot du texte "sa forme normalement correcte ou une forme qui lui est flexionnellement liée" (Baranes & Sagot 2014: 137). Ainsi, les données textuelles normalisées dans le cadre de l'analyse quantitative (voir section 4.1) ont été utilisées à nouveau. Puisque les classifieurs d'apprentissage supervisé reposent sur la représentation numérique, les données textuelles doivent être transformées en vecteurs. Pour ce faire, nous avons exploité la méthode du sac de mots (*bag of words* ou BOW ; Salton & Buckley 1988) et celle de TF-IDF. La première consiste à représenter chaque document par les vecteurs de fréquence des mots qui le composent, permettant d'apercevoir la fréquence de chaque mot au sein de l'ensemble de données. La méthode de TF-IDF, une approche avancée, consiste à pondérer les mots distinctement plus fréquents dans un document spécifique que dans d'autres documents. Plus précisément, elle attribue un poids élevé aux mots qui apparaissent fréquemment dans un document particulier mais pas dans de

nombreux documents du corpus. Les mots-outils, par exemples, sont peu significatifs et leur importance est sous-pondérée grâce à cette méthode.

En complément des données textuelles, les caractéristiques linguistiques observées et retenues lors de l'analyse du corpus, transformées en vecteurs, ont été intégrées à l'expérience. Les traits suivants ont été extraits, de manière automatique: les catégories morphosyntaxiques disponibles sur l'étiqueteur StanfordCoreNLP (nom, déterminant, adjectif, etc.), les diverses formes verbales (futur, conditionnel, préfixe *re-*, etc.), les modifieurs (*ne*, *jamais*, *aucun*, *mais*, etc.), la polarité des phrases et des mots obtenus selon la bibliothèque TextBlob¹⁶, la position de la phrase dans un avis¹⁷ et la forme de surface¹⁸.

L'apprentissage supervisé de surface s'appuie sur ces traits qui permettent de détecter automatiquement différents types d'évaluation. Le classifieur effectue une prédiction parmi les six possibilités: opinion positive/négative/mixte, suggestion, intention et description. Afin d'évaluer la capacité de généralisation d'un modèle, c'est-à-dire de demander aux modèles de refaire leurs prédictions sur de nouvelles données jamais traitées, une validation croisée stratifiée à cinq plis a été effectuée. La meilleure F-mesure a été obtenue (0,88) grâce au classifieur SVM linéaire. De plus, les performances pour chaque catégorie sont: opinion positive (0,94), opinion négative (0,72), opinion mixte (0,69), suggestion (0,84), intention (0,88), et description (0,41). L'opinion positive est détectée avec la meilleure performance, suivie par l'intention et la suggestion. En revanche, l'opinion négative donne un score inférieur aux précédents. La détection de la description connaît la plus mauvaise performance. Ce résultat est dû à la faible fréquence de la description dans le corpus. Cette catégorie semble donc s'appuyer sur peu de marqueurs lexicaux en raison de la nature hétérogène des informations qui y sont présentes. Par conséquent, la description a tendance à être rangée dans la classe majoritaire, c'est-à-dire l'opinion positive, et occasionnellement dans l'opinion négative. L'opinion mixte semble également engendrer des difficultés aux classifieurs.

6. Conclusion

Le travail présenté dans cet article porte sur l'analyse d'évaluations de restaurants publiées en ligne par des clients. Il montre que les évaluations en ligne peuvent avoir plusieurs fonctions et contenir différentes informations: le

¹⁶ Plus précisément, le score de polarité, le score de subjectivité, la fréquence des mots positifs et négatifs obtenus grâce à l'utilisation de la bibliothèque TextBlob (<https://github.com/sloria/textblob-fr> [consulté le 25 octobre 2021]) et les émoticônes. TextBlob est l'une des bibliothèques les plus utilisées en TAL.

¹⁷ Il s'agit d'une variable catégorielle qui prend pour valeur F (première phrase), L (dernière phrase), M (phrase au milieu) ou S (un avis composé d'une seule phrase).

¹⁸ Le nombre de caractères, la longueur de la phrase, la diversité et la densité des mots, les ponctuations multiples (!!!) et les majuscules en cascade (EXCELLENT), ainsi que l'unité monétaire (euro et €).

client laisse son évaluation afin de donner son avis (opinion), faire des suggestions d'amélioration (suggestion), exprimer ses intentions (intention) ou décrire son expérience (description). Le corpus annoté avec ces informations a servi de référence pour effectuer différentes analyses quantitatives grâce aux méthodes de la linguistique de corpus anglo-saxonne. Chaque catégorie a été étudiée et décrite à l'aide de mots-clés, de la distribution des parties du discours et de sa localisation au sein d'un avis. L'opinion se caractérise ainsi par un lexique évaluatif et les cibles sur lesquelles elle porte et se manifeste au moyen de syntagmes nominaux. La suggestion, l'intention et la description comportent des syntagmes verbaux qui permettent au client de suggérer une action future au restaurant ou aux clients potentiels (la suggestion), d'effectuer une action (l'intention) ou de décrire l'arrière-plan de l'expérience (la description). Plus précisément, la description est liée, d'une part, aux informations factuelles ayant un impact avant la visite du restaurant, telles que le motif de cette dernière ou la précision sur une réservation effectuée. D'autre part, elle est relative à des informations factuelles concernant l'expérience vécue lors de la visite, comme le choix du menu ou son prix. En ce qui concerne la localisation des catégories dans un avis, l'opinion et la description sont souvent placées au début d'un avis, alors que la suggestion et l'intention sont positionnées à la fin. En nous appuyant sur l'observation manuelle, nous avons également développé un module de détection automatique de chacune de ces fonctions d'évaluation. Le résultat montre, à ce stade, des performances inégales: les opinions positives, intention et suggestions tendent à être correctement identifiées, contrairement aux opinions mixtes et description. Le couplage des perspectives et techniques de la linguistique de corpus et du TAL a permis de mettre en évidence des propriétés du corpus qu'un seul angle n'aurait pas permis d'observer. Les méthodes utilisées pour l'analyse peuvent être généralisées et reproduites sur d'autres données.

BIBLIOGRAPHIE

- Amiot, D. (2012): De la localisation à l'évaluation: des verbes préfixés évaluatifs au sens bien particulier, *Corela* [En ligne], HS-12 | 2012. Disponible: <http://journals.openedition.org/corela/2775>.
- Amiot, D. & Stosic D. (2015): Morphologie aspectuelle et évaluative en français et en serbe. *Lexique*, 22, 111-142.
- Anthony, L. (2019): *AntConc* (version 3.5. 8)[computer software]. Tokyo (Waseda University).
- Baccini, A. (2010): *Statistique Descriptive Multidimensionnelle (pour les nuls)*. Institut de Mathématiques de Toulouse-UMR CNRS, 5219.
- Baker, P. (2010): *Sociolinguistics and corpus linguistics*. Edinburgh (Edinburgh University Press).
- Baranes, M. & Sagot B. (2014), Normalisation de textes par analogie: le cas des mots inconnus. Actes de la 21ème conférence sur le Traitement Automatique des Langues Naturelles, 137-148.
- Benamara, F., Asher N., Mathieu Y.-Y., Popescu V. & Chardon B. (2016): Evaluation in discourse: A corpus-based study. *Dialogue & Discourse*, 7(1), 1-49.
- Benzécri, J.-P. (1969): Statistical analysis as a tool to make patterns emerge from clouds. In S. Watanabe (éd.), *Methodology of Pattern Recognition*. New York (Academic Press), 35-74.
- Biber, D. & Finegan E. (1988): Adverbial stance types in English. *Discourse processes*, 11(1), 1-34.



- Bifet, A., Holmes, G. & Pfahringer, B. (2011). Moa-tweetreader: real-time analysis in twitter streaming data. International conference on discovery science. Berlin (Springer), 46-60.
- Bratman, M. (1987): Intention, Plans, and Practical Reason. Cambridge (Harvard University Press).
- Brun, C. & Hagège C. (2013): Suggestion Mining: Detecting Suggestions for Improvement in Users' Comments. Research in Computing Science, 70, 199-209.
- Carreiras, M., Gernsbacher M.-A. & Villa V. (1995): The advantage of first mention in Spanish. Psychonomic Bulletin & Review, 2(1), 124-129.
- Charaudeau, P. (1992): Grammaire du sens et de l'expression. Vanves (Hachette Éducation).
- Charest, S., Brunelle, É., Fontaine, J. & Pelletier, B. (2007): Élaboration automatique d'un dictionnaire de cooccurrences grand public. Actes de la 14ème conférence sur le Traitement Automatique des Langues Naturelles, 265-274.
- Cheminée, P. (2006): 'Vous avez dit clair'? Le lexique des pianistes, entre sens commun et terminologie. Cahiers du LCPE: Dénomination, désignation et catégories, 7, 39-54.
- Desagulier, G. (2020): Multivariate Exploratory Approaches. In M. Paquot & S. Th. Gries (éds.), Practical handbook of corpus linguistics. Berlin & New York (Springer).
- Dunning, T. E. (1993): Accurate methods for the statistics of surprise and coincidence. Computational linguistics, 19(1), 61-74.
- Eshkol-Taravella, I. & Kang, H.-J. (2019): Observation de l'expérience client dans les restaurants. Actes de la 26ème conférence sur le Traitement Automatique des Langues Naturelles, 361-370.
- Fleiss, J.-L., Levin, B. & Paik, M.-C. (1981): The measurement of interrater agreement. Statistical methods for rates and proportions, 2(212-236), 22-23.
- Fradin, B. (2003): Le traitement de la suffixation *en-et*. Langues, 51-77.
- Ganu, G., Elhadad N. & Marian A. (2009), Beyond the Stars: Improving Rating Predictions using Review Text Content. In 12th International Workshop on the Web and Databases.
- Gernsbacher, M.-A. (1990), Language Comprehension as Structure Building, London (Psychology Press).
- Goldberg, A.-B., Fillmore N., Andrzejewski D., Xu Z., Gibson B. & Zhu X. (2009): May All Your Wishes Come True: A Study of Wishes and How to Recognize Them. In Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, 263-271.
- Gosselin, L. (2010): La validation des représentations. Les modalités en français. Leiden (Brill).
- Gosselin, L. (2017): Les modalités appréciatives et axiologiques. Sémantique des jugements de valeur. Cahiers de lexicologie: Revue internationale de lexicologie et lexicographie, 111, 97-119.
- Gower, J. C. (1968): Adding a point to vector diagrams in multivariate analysis. Biometrika, 55(3), 582-585.
- Haberman, S. J. (1973): The analysis of residuals in cross-classified tables. Biometrics, 205-220.
- Hatzivassiloglou, V. & McKeown, K. (1997): Predicting the semantic orientation of adjectives. In 35th annual meeting of the association for computational linguistics and 8th conference of the european chapter of the association for computational linguistics, 174-181.
- Hunston, S. & Thompson, G. (2000): Evaluation in text: Authorial stance and the construction of discourse: Authorial stance and the construction of discourse. Oxford (Oxford University Press).
- Kang, H.-J. & Eshkol-Taravella, I. (2020): Les avis sur les restaurants à l'épreuve de l'apprentissage automatique. Actes de la 27ème conférence sur le Traitement Automatique des Langues Naturelles, 249-257.
- Kim, S.-I., Lee J.-H. & Gernsbacher M.-A. (2004): The advantage of first mention in Korean the temporal contributions of syntactic, semantic, and pragmatic factors. Journal of psycholinguistic research, 33(6), 475-491.
- Landis, J.-R. & Koch G.-G. (1977): The Measurement of Observer Agreement for Categorical Data. Biometrics, 33(1), 159-174.
- Lê, S., Josse J. & Husson F. (2008): FactoMineR: an R package for multivariate analysis. Journal of statistical software, 25(1), 1-18.

- Lebart, L., Pincemin B. & Poudat C. (2019), Analyse des données textuelles, Mesure et évaluation. Québec (Presses de l'Université du Québec), t. 11.
- Lefer, M.-A. & Grabar N. (2013): French evaluative prefixes in translation: from automatic alignment to semantic categorization. In proceedings of Theoretical and computational morphology: New trends and synergies workshop, International Congress of Linguists.
- Lefer, M.-A. & Grabar N. (2014): Evaluative prefixes in translation: From automatic alignment to semantic categorization. *Linguistic Issues in Language Technology*, 11(6), 169-187.
- Legallois, D. & Ferrari S. (2006): Vers une grammaire de l'évaluation des objets culturels. Actes d'ISDD06: Colloque international Discours et Document, Schedæ, 57-68.
- Liu, B. (2015): Opinions, Sentiment, and Emotion in Text, *Sentiment Analysis: Mining Opinions, Sentiments, and Emotions*. Cambridge (Cambridge University Press).
- Martin, J.-R. & White P.-R. (2003): *The language of evaluation*. New York (Springer).
- Mathieu, Y.-Y. (2006): A computational semantic lexicon of french verbs of emotion. *Computing Attitude and Affect in text: Theory and Applications*. New York (Springer), 109-124.
- McEnery, T. (2016): Keywords. In P Baker & J. Egbert (eds.), *Triangulating Methodological Approaches in Corpus Linguistic Research*. London (Routledge), 32-44.
- Negi, S. & Buitelaar P. (2015): Towards the Extraction of Customer-to-Customer Suggestions from Reviews. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, 2159-2167.
- Negi, S., Daudert T. & Buitelaar P. (2019): SemEval-2019 Task 9: Suggestion Mining from Online Reviews and Forums In *Proceedings of the 13th International Workshop on Semantic Evaluation*, 877-887.
- Polanyi, L. & Annie Zaenen (2006): Contextual valence shifters. In J.G. Shanahan, Y. Qu & J. Wiebe (eds.), *Computing attitude and affect in text: Theory and applications*. New York (Springer), 1-10.
- Poole, R. (2016): Good times, bad times: A keyword analysis of letters to shareholders of two Fortune 500 banking institutions. *International Journal of Business Communication*, 53(1), 55-73.
- Poudat, C. & Landragin F. (2017): *Explorer un corpus textuel: Méthodes-pratiques-outils*. Berlin (De Boeck Supérieur), Édition Kindle.
- Poudat, C., Wigham C.-R. & Liégeois L. (2020): *Les corpus de la communication médiée par les réseaux: une introduction, corpus, corpus complexes. Traitements, standardisation et analyse des corpus de communication médiée par les réseaux*, 20, Édition Kindle.
- Pupier, P. (1998): Une première systématique des évaluatifs en français. *Revue québécoise de linguistique*, 26(1), 51-78.
- Pustejovsky, J. & Stubbs, A. (2012). *Natural Language Annotation for Machine Learning: A guide to corpus-building for applications*. Sebastopol (O'Reilly Media, Incorporated).
- Ramanand, J., Bhavsar K. & Pedanekar N. (2010): Wishful Thinking - Finding suggestions and 'buy' wishes from product reviews. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, 54-61.
- Salton, G. & Buckley C. (1988): Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, 24(5), 513-523.
- Scott, M. & Tribble C. (2006): *Textual patterns: Key words and corpus analysis in language education*. Amsterdam (John Benjamins Publishing).
- Seale, C., Ziebland S. & Charteris-Black J. (2006): Gender, cancer experience and internet use: a comparative keyword analysis of interviews and online cancer support groups. *Social science & medicine*, 62(10), 2577-2590.
- Taboada, M., & Trnavac, R. (2013): *Nonveridicality and evaluation: Theoretical, computational and corpus approaches*. Leiden (Brill).
- Vasquez, C. (2014): *The discourse of online consumer reviews*. London & New York (Bloomsbury Publishing), Edition Kindle.
- Vernier, M. & Monceaux L. (2010): Enrichissement d'un lexique de termes subjectifs à partir de tests sémantiques. *Traitement Automatique des Langues*, 51(1), 125-149.

- Wicaksono, A.-F. & Myaeng S.-H. (2012): Mining Advices from Weblogs. In Proceedings of the 21st ACM International Conference on Information and Knowledge Management, 2347-2350.
- Wicaksono, A.-F. & Myaeng S.-H. (2013): Automatic Extraction of Advice-Revealing Sentences Foradvice Mining from Online Forums. In Proceedings of the 7th International Conference on Knowledge Capture, 97-104.
- Zhang, L. (2012): Analyse automatique d'opinion: problématique de l'intensité et de la négation pour l'application à un corpus journalistique, thèse de doctorat. Caen (Université de Caen).