Qu'en pensent les Chinois ?

Essai d'exploration de l'opinion publique chinoise à travers des documents disponibles sur la toile.

[Bad karma]

Liangcai Shen, André Salem[1]
liangcaishen@gmail.com, salem@msh-paris.fr

Résumé : Les nombreux moyens d'expressions liés aux technologies du web deviennent chaque jour plus accessibles aux citoyens chinois désireux d'exprimer leurs réactions à propos de sujets d'actualité. A propos d'un incident médiatique entraîné par les propos d'une célèbre actrice américaine, après une catastrophe naturelle survenue en Chine, nous avons cherché à mettre à jour quelques-unes des dimensions de la réaction suscitée par ces propos dans l'opinion publique chinoise. Pour cette première étude, nous avons choisi de comparer quelques échantillons de textes publiés sur la toile par la presse officielle, des textes relevés sur des blogs personnels et des interventions collectées sur des forums publics. Cette première démarche, aux dimensions modestes, illustre la possibilité et l'intérêt du type d'enquête proposé.

Mots-clés : Etude d'opinion, médias, textométrie

Abstract : The expressions of many ways related to web technologies become ever more accessible to Chinese citizens wishing to express their opinions about topical issues. About an incident caused by the remarks of an american actress, after a natural disaster in China, we try to explore the dimensions of the reaction to these words in chinese public. We chosed to compare a few samples of texts published on the web by the official press, texts recorded on personal blogs and responses collected from public forums. This first approach illustrates the ability and interest of the type of proposed investigation.

Keywords : Opinion studies, media analysis, textometrics,

摘要：如今的网络科技五花八门，一日千里，中国老百姓现在可以更轻松自如，随心所欲地表达自己对时事热点的各种意见和看法。中国汶川重大自然灾害之后，美国某影星的某些个人看法在华人世界中激起了轩然大波。为此，作者分别选取了公布于官方新闻网站上的新闻与评论，记录在个人博客中的文章以及公众论坛中的回帖等内容作为比较样本，并运用词量法来分析和对比中国社会各阶层民众的语料语库，进而实现了跟进该事件在国民大众舆论中不同反响，为民众观点探究开拓了新的方法。此端倪之作，文集容量虽小，但该方法却充分证明了它在社会舆论调查上的新潜力，并且也拓展中文自然语言处理的新领域。

关键词：词量法，媒体，观点探究，民众意见，百姓观点，抵制

Comment cerner les sentiments suscités dans les différentes couches de l'opinion publique chinoise par un événement médiatique dont la répercussion a été planétaire, compte tenu des moyens modernes de circulation de l'information ? S'agissant d'un pays aussi étendu et aussi diversifié que la Chine, la question peut paraître naïve voire dénuée de sens. Le titre quelque peu provocateur que nous avons donné à cette étude souligne en fait le caractère chimérique d'une telle entreprise conçue comme une tentative d'exploration exhaustive aboutissant à des conclusions nettes et clairement formulées.

Cependant, par-delà la multiplicité et la diversité des réactions individuelles susceptibles d'être observées au sein d'un peuple qui compte plus d'un milliard d'habitants, dont les

[1] Les auteurs remercient Jean-Maxence Granier, de la société *Think-Out*, pour ses précieuses suggestions et ses encouragements dans la réalisation de cette étude.

langues, les coutumes, sont extrêmement variées, les technologies liées au web offrent désormais la possibilité d'observer, et ce quel que soit l'endroit où l'on se trouve sur la planète, des interactions entre citoyens chinois qui échangent des avis par ce biais. Il est bien entendu que cet échantillon de réactions, limité aux intervenants qui ont accès à ces nouveaux médias, ne constitue en aucun cas une photographie fidèle de l'ensemble de la société chinoise moderne. Cependant, de part le fait qu'elles aient été produites par des acteurs de la vie économique chinoise, ces réactions constituent un matériau extrêmement précieux pour ceux qui s'intéressent à l'étude de l'opinion publique chinoise.

Dans ce qui suit, nous commencerons par rappeler le contexte des évènements qui ont été au centre de notre enquête (§1). Nous présenterons ensuite les différents supports que notre stratégie de fouille de textes à permis de repérer en liaison avec notre recherche (§2). La section suivante (§3) est consacrée au dépouillement du corpus. La dernière section (§4) analyse les emplois contextuels de la forme *boycott,* particulièrement fréquente dans les forums.

1 Contexte de la recherche

S'exprimant à l'occasion du festival de cinéma de Cannes, le 24 mai 2008, quelques jours après le tremblement de terre survenue en Chine[2], l'actrice américaine Sharon Stone a tenté d'établir un parallèle entre cette catastrophe naturelle et l'action politique de l'État chinois au Tibet, possiblement responsable, selon elle d'un, d'une altération du *karma* commun aux chinois. La connotation particulièrement sensible de ce type de conclusions dans le monde sinophone, qui implique, d'une certaine manière, une punition *méritée* par ses victimes, a immédiatement suscité de très vives réactions dans l'opinion publique chinoise.

Dans cette étude nous avons tenté une première exploration des réactions à cet événement à partir des textes accessibles sur l'Internet. Dans un premier temps, nous avons interrogé deux moteurs de recherche (***Google*** et ***Baidu***) pour localiser les documents qui contenaient à la fois les termes : 莎朗-斯通 Shalang-Sitong (Sharon Stone en caractères chinois), et 四川 (Sichuan, région dans laquelle a eu lieu le tremblement de terre 汶川 Wenchuan). Au vu des résultats, nous avons constitué un premier corpus à partir de trois types de sources textuelles différentes que nous avions identifiées :

- des sites d'information en ligne (presse, agences, etc.) ;
- des textes présentés sur des blogs personnels ;
- des réactions individuelles sur des forums ouverts au public par des sites très fréquentés.

Bien entendu, dans cette première étude, qui ne porte que sur un corpus restreint, nous ne prétendrons pas rendre compte de la totalité des réactions suscitées à cette occasion dans l'ensemble de la population chinoise. Il nous semble cependant qu'elle montre la possibilité de recueillir et de classer une certaine variété de réactions qui diffèrent largement selon le media utilisé et dont certaines présentent une fréquence importante au sein d'un même média. Nous nous proposons essentiellement de montrer la possibilité de réaliser de manière relativement simple, une enquête de ce type.

[2] Une série de tremblements de terre survenus autour du 12 mai 2008, dans la région de Si Chuan (Chine), a causé la mort de plus de 69 000 personnes et entraîné de très importantes destructions dans toute la région. Ces circonstances ont été à l'origine d'une campagne nationale et internationale de solidarité avec les victimes.

Points de repères

- *Les propos incriminés (Cannes 24 mai 2008)*[3]

  And I have been concerned about... oh… how should we deal with the Olympics because they haven't been nice to the Dalai Lama, who is a good friend of mine.. and the earthquake…and all the stuff happened… I think … is that karma ? when you're not nice that the bad things happen.

- *Le karma*[4] *:*

  (sanskrit कर्म, de la racine *kri*, *acte*, *action*) est un terme utilisé dans plusieurs religions orientales. Le karma désigne le cycle des causes et des conséquences lié à l'existence des êtres sensibles. Le *karma* est la somme de ce qu'un individu a fait, est en train de faire ou fera. Dans les religions incorporant les concepts de réincarnation, les effets de ces actes *karmiques* se répercutent sur les différentes vies d'un individu. Chaque être y est responsable de son karma et donc de sa sortie du Samsara.

- *La notion de* ***« karma négatif »*** *:*

  La traduction chinoise diffusée dans les médias à partir de cette déclaration initiale : 报应 *bao ying* . (*karma négatif),* accentue peut-être, en les synthétisant, le caractère blessant des propos tenus par l'actrice. L'expression possède une connotation particulièrement négative dans le monde sinophone de *mauvaise conséquence justement méritée (punition méritée)*.

Chronologie sommaire

avril2008 : agitation suivie d'une répression dans la province chinoise du Tibet.
incidents et manifestations contre le gouvernement chinois sur le parcours de la flamme olympique dans plusieurs pays occidentaux.
tremblement de terre dans la province de Si Chuan (70 000 victimes)

mai 2008 : déclarations de Sharon Stone au festival de Cannes (cf. supra)
Vives réactions dans la presse, sur les blogs et les forums chinois
déclaration de Dior Chine se désolidarisant de l'opinion de Sharon Stone

juin 2008 : excuses officielles de Sharon Stone
la campagne de réactions se poursuit sur les forums

août 2008 : ouverture des jeux olympiques à Pékin

---

[3] Propos retranscrits d'après la vidéo enregistrée pendant l'interview et postée, entre autres, sur le site http://lionelshen.free.fr/Labo/stage/ST_Karma_interview.flv

[4] Les auteurs se sont largement inspirés de l'article *karma* de l'encyclopédie en ligne *Wikipédia*.

2 Localisation et présélection des textes

Dans un premier temps, nous avons entrepris de localiser, sur la toile, les textes susceptibles de concerner le débat créé parmi les internautes chinois. Nous avons utilisé de manière complémentaire deux moteurs de recherche : le classique ***Google***, dans sa version www.google.cn[5], mais également le moteur ***Baidu*** 百度[6], réputé plus performant pour la recherche des documents numériques rédigés en chinois.

Figure 1a

Recherche sur ***Google.cn*** à partir du mot-clé (Sharon Stone)

a) en haut les suggestions du moteur de recherche
b) les premiers résultats référencés

---

[5] Signalons que la version chinoise du moteur www.google.cn est placée sous le contrôle effectif des autorités chinoises.

[6] Le moteur de recherche ***Baidu*** a été créé par des chercheurs sinophones expatriés aux Etats-Unis. En Chine continentale, sa popularité dépasse largement celle de son concurrent *Google*.

Les réponses fournies par les deux moteurs présentaient une grande intersection pour ce qui concerne les sites officiels (sites de presse, etc.). Comme prévu, la couverture du moteur ***Baidu***, s'est révélée plus importante pour ce qui concerne les blogs et les forums.

Le moteur ***Google*** nous fournit de très nombreuses références concernant ce débat à partir du seul mot-clé *Sharon Stone*. La recherche promptée par ***Google***, nous propose de choisir entre :

- Sharon Stone 1 160 000 résultats indexés, 1ère suggestion
- *Sharon Stone et ses films* 15 600 résultats indexés, 2ème suggestion
- *Sharon Stone, le tremblement de terre à Sichuan cause du karma négatif des chinois* 78 700 résultats indexés, 3ème suggestion

C'est cette dernière suggestion qui correspond manifestement le mieux à notre recherche.

Figure 1b

Recherche sur le moteur ***Baidu*** à partir des mots-clé (*Sharon Stone* et *Sichuan*)
(premiers résultats référencés)

À partir des deux mots-clé : *Sharon Stone* et *Sichuan*, le moteur ***Baidu*** nous fournit une liste de références qui, si elle recoupe largement la précédente liste fournie par ***Google.cn*** en ce qui concerne les sites officiels, est beaucoup plus abondante pour ce qui concerne les sites de type « forum ».

Parmi les trois types de médias qui apparaissent dans les références (presse en ligne, blogs, forums) nous avons choisi, pour cette étude exploratoire, de sélectionner un échantillon de sites plus particulièrement référencés par les moteurs de recherche[7].

---

[7] Le calcul de l'indice de référencement (*ranking*) qui sert à trier les sites dans les résultats d'un moteur de recherche s'appuie en principe sur le nombre des consultations effectuées sur chacun des sites.

Nous avons choisi de retenir pour notre étude les textes publiés sur l'Internet entre le 26 mai 2008 (date des premières réactions) et le 02 octobre 2008 (clôture du fil de discussion sur ce sujet sur les forums observés). Nous avons privilégié les sites les plus fréquentés par les internautes à partir des indices de fréquentation calculés par les différents moteurs de recherche (ranking).

2.1 La presse en ligne

Comme partout ailleurs, les grands quotidiens nationaux chinois et les agences de presse entretiennent des sites informationnels sur l'Internet qui leur permettent de mettre leurs principales publications à la disposition des internautes dans des délais relativement courts. Nous avons sélectionné vingt-sept mis en ligne après leur parution par ces grands organes de la presse officielle.

Le site de l'agence 新浪 Sina (Nouvelle vague) et celui de l'agence 新华社 Xinhua (Chine nouvelle) ont constitué nos principales sources lors de cette sélection de notre volet de *Presse en ligne.* À partir des sites sélectionnés par les deux moteurs de recherche utilisés, nous avons retenu 27 articles signalés comme ayant été le plus souvent consultés par les internautes.

2.2 Les blogs

Dans le contexte chinois comme dans le contexte francophone, le concept *blog* peut recouvrir des situations très différentes : un journal intime assumé ou anonyme, un journal d'opinion tenu par un journaliste, les échanges quotidiens d'une classe de collège, une œuvre littéraire collective en construction, etc. Comme partout dans le monde, le phénomène connaît en Chine un immense succès grâce à une grande facilité de publication en ligne, une relative tolérance éditoriale et une grande capacité d'interaction avec le lectorat.

Le blog est en général édité et mis à jour par un auteur ou un groupe d'individus identifiés qui ne donnent que très rarement aux lecteurs potentiels la possibilité de s'exprimer à leur tour sur le site du blog. Le nombre de lecteurs d'un blog surpasse souvent celui des lecteurs d'une publication traditionnelle sur papier. En très peu de temps, certains blogs sont devenus extrêmement fréquentés[8] au sein de la communauté des internautes chinois en Chine et à l'étranger).

Nous avons sélectionné vingt-six blogs parmi les plus fréquentés, nous avons veillé à rassembler des opinions différentes autant qu'il se pouvait. Ces blogs nous ont fourni un matériau à peu près comparable (du point de vue du volume de texte) à la partie sélectionnée pour représenter la presse dans notre corpus.

Signalons enfin que les moteurs de recherche proposent des outils spécifiques pour la recherche des blogs (*blogsearch.google.cn, blogsearch.baidu.com)* qui nous ont permis de localiser ces derniers sans difficulté.

[8] On trouvera plus loin l'exemple d'un blog consulté par plus d'un million d'utilisateurs au cours de la semaine considérée.

sina 影音娱乐 | 影音娱乐 > 明星全接触 > 正文

2

Dior迪奥声明：拒不认同莎朗-斯通言论

http://www.sina.com.cn 2008年05月27日16:10 新浪娱乐

莎朗-斯通对四川地震发冷血言论 引起各方声讨

3

视频：莎朗-斯通对于四川地震的冷血言论

新浪娱乐讯 今日（2008年5月27日）下午3点15分，法国迪奥中国区公关负责人致电新浪，就迪奥化妆品代言人莎朗-斯通“四川地震有趣”等恶性言论表态，要求通过新浪网发布声明。

声明如下：

4

美国好莱坞影星莎朗-斯通于2008年5月24日在法国戛纳电影节接受香港有线电视采访时，就中国四川汶川地震发表了她的个人言论，我们对此未经深思的突发言论绝不认同，也深表遗憾。

Dior迪奥是最早进入中国的国际品牌之一，深受广大消费者的喜爱和尊重，我们决不支持任何伤害中国人民情感的言论。

我们对此次四川汶川大地震中不幸遇难的同胞表示哀悼，并对灾区的人民表示深切的同情和慰问。我们重申对中国市场的长期承诺并对灾区重建予以鼎力支持。

5

特此声明！

Dior迪奥-中国

二零零八年五月二十六日

声明：新浪网登载此文出于传递更多信息之目的，并不意味着赞同其观点或证实其描述。

6

【莎朗-斯通吧】

【发表评论 1847条】

7

网页 新闻 搜索 Powered By Google 【参加“我的2008”，赢取2008万大奖】

Figure 2

Volet *Presse* du corpus ***StoneKarma***
Déclaration de *Dior-Chine* reproduite par l'agence *Sina*
Retranscription intégrale de la déclaration de Dior-Chine

Figure 3

Volet ***Blogs*** du corpus ***StoneKarma***
Les propos de l’actrice reproduits sur le blog de HU Jianli

14 15 16 17 18 19 20

網易 NETEASE www.163.com 网易跟帖 网易论坛 > 社会跟贴6版 评论共16069条 显示12181条

话题：美影星莎朗斯通称四川地震是报应引声讨(视频) 查看原文

躲得过的怪物，躲不过的刺激 边看新闻边游戏，轻送玩转大富翁 悄悄在MM之间流行的游戏

» 热门评论 1 2 3 4 5 ··· 下一页

网易湖北武汉网友 ip：58.49.*.*： 2008-05-27 03:10:08 发表

私通，肯定是他妈与他爷爷私通的

回复 收藏到博客 支持[1048] 反对[147] 举报

网易河南郑州网友 ip：222.143.*.*： 2008-05-27 03:[illegible] 发表

我看这个胸大无脑的女人迟早也会遭到报应的！！！！

回复 收藏到博客 支持[789] 反对[88] 举报

网易山东东营网友 ip：60.214.*.*： 2008-05-27 03:45:08 发表

网易加拿大网友(70.69.*.*) 的原贴：

大家团结起来，让她知道中国人民的力量！坚决抵制法国迪奥化妆品(莎朗斯通代言),莎朗斯通的影片！！

网易辽宁大连网友(116.3.*.*) 的原贴：

大家团结起来，让她知道中国人民的力量！坚决抵制法国迪奥化妆品(莎朗斯通代言),莎朗斯通的影片！！

大家团结起来，让她知道中国人民的力量！坚决抵制法国迪奥化妆品(莎朗斯通代言),莎朗斯通的影片！！

回复 收藏到博客 支持[768] 反对[101] 举报

网易北京昌平网友 ip：221.218.*.*： 2008-05-27 03:28:01 发表

想在中国有一席之地的世界各种品牌，以后请不要再用沙朗.斯通做代言人了，否则可能会在中国遇到很尴尬的处境的

回复 收藏到博客 支持[645] 反对[59] 举报

网易广东东莞网友 ip：125.93.*.*： 2008-05-27 03:49:25 发表

年近五十还要卖肉,你也真是报应啊!

回复 收藏到博客 支持[565] 反对[61] 举报

网易福建福州网友 ip：59.61.*.*： 2008-05-27 03:49:31 发表

迪奥,DIOR
如果不更换代言就离开中国吧.

回复 收藏到博客 支持[537] 反对[46] 举报

Traduction de la dernière intervention (IP 59.61.*.*) :
Dior, si tu ne changes pas de représentante, quitte la Chine !

Figure 4

Volet *Forum* du corpus ***StoneKarma***
Exemples de discussion sur le forum *NetEase*

Guide de lecture pour les figures 3 à 5

Principaux composants du site de l'agence *Sina*

1. Nom du site Internet (Sina)
2. Titre de l'article (Déclaration de Dior Chine)
3. Hyperlien de la Vidéo de l'interview de Sharon Stone
4. Le corps de la déclaration de Dior se désolidarisant de l'actrice
5. Datation et signature de Dior Chine
6. Proposition du thème de la discussion
7. Nombre de réactions (1847 fois à ce jour)

Principaux composants du blog de HU Jianli

(1 million de visiteurs dans cette période)

8. Nom de l'auteur du blog
9. De haut en bas : 1) Notation du blog numéro 1 (fréquentation globale) [9];
2) Notation du blog numéro 2 (mise à disposition et réactions),
3) Nombre de visites (1 067 602 dans la période)
10. Vidéo de l'interview de Sharon Stone
11. Retranscription des propos incriminés
12. Déclaration de Dior se désolidarisant de l'actrice
13. Commentaires de l'auteur du blog

Principaux composants du forum du site de *Netease*

14. Nom du forum du site (Netease – www.163.com)
15. Thème de la discussion (Les propos de ST sur le séisme à Sichuan provoquent des appels à sanctions. Avec des extraits vidéos)
16. Nombre de réponses (16 069)
17. Numéro IP de l'internaute (partiel, ne permettant pas de l'identifier totalement, signalant cependant que le site est capable de localiser l'émetteur)
18. Nombre de ceux qui ont voté « pour » cette proposition (ici : 1048 pour)
19. Nombre de ceux qui ont voté « contre » cette proposition (ici : 147 contre)
20. Citation de propos d'un internaute s'étant exprimé précédemment, reprise dans la réaction d'un internaute suivant.

---

[9] Cette note de *popularité* est attribuée aux blogs individuels par le gestionnaire du site *Sina* en fonction du nombre de visites reçues par chacun d'eux.

2.3 Les forums

Un *forum en ligne* est un site d'échanges entre internautes se situant au même niveau du point de vue éditorial. Les discussions y prennent place sous la forme de « fils » de messages, publication instantanée ou différée ; cette publication est souvent durable, car les messages ne sont pas effacés. Elle est par nature le fait de plusieurs auteurs. Dans certains forums à inscription, les messages sont modifiables a posteriori par leurs auteurs.

Les fonctionnalités offertes par les différents forums (citation d'un point de vue précédemment exprimé, création d'intertitres, mise en page/indentation particulière, modération des droits d'accès, a priori ou a posteriori...) peuvent varier d'un forum à l'autre : certains forums ne permettent que de contribuer de manière ponctuelle à un sujet discussion, tandis que d'autres permettent de répondre plus longuement à un message particulier, voire à un paragraphe particulier contenu dans ce message.

Pour représenter le type de média *forum*, nous avons sélectionné l'ensemble du fil de discussion *: L'actrice Sharon Stone considère que le tremblement de terre de Sichuan est une conséquence d'un mauvais karma, ce qui entraîne de vives réactions* qui se sont développé sur le site de l'agence NetEase (www.163.com) [10]. Nous avons choisi ce site, parmi d'autres parce qu'en dépit d'un thème nominal au caractère réprobateur, il présentait, à première vue, une discussion beaucoup plus ouverte à des opinions variées que des sites concurrents. Comme on le verra plus loin, l'intégrale de la discussion, au sein de laquelle nous nous sommes refusés à faire des sélections, présentait un volume beaucoup plus important que les deux autres volets du corpus.

Nous appellerons désormais ***StoneKarma*** le corpus ainsi rassemblé. Ces trois volets, prélevés sur des supports électroniques de différents types, englobent *grosso modo* trois sources qui peuvent prétendre représenter en partie l'opinion publique chinoise. Le choix de ces différents supports permet de mieux cerner l'hétérogénéité de cette opinion publique qui s'exprime sur le web. On peut supposer, a priori, que la presse représente, dans la plupart des cas, l'opinion officielle des autorités chinoises. Les blogs et les forums fournissant une approche moins contrôlée de l'opinion des citoyens.

3 Dépouillement quantitatif du corpus

L'ensemble du corpus compte 512 806 caractères chinois (balises comprises) que le segmenteur isole en 208 707 occurrences de mots chinois[11]. On peut diviser le corpus en 16 953 paragraphes. Ces paragraphes correspondent à des retours à la ligne dans les textes de presse et dans les blogs et à des successions de tours de paroles.

Les textes rassemblés dans le volet *Presse* sont au nombre de 27, les textes de blogs au nombre de 26 et le volet forum est constitué de 3 023 interventions individuelles.

---

[10] Le nom du site : ***163*** constitue un jeu de mots à partir de la forme phonétique de l'expression « tout va bien /ou avoir le vent en poupe » (一路平安, yi lu pin an).

[11] La notion de « mot chinois » et la segmentation automatique en « mots » seront précisés plus loin (§2).

Tableau 1
Exemple de réactions dans la partie *forum* du corpus

Un forum : Le forum du site Netease (163.com)

<media=forum>
网易 论坛 话题 ： 美 影星 莎朗斯 通称 四川 地震 是 报应 引 声讨
评论
网易 广西 桂林 网友 [moqingli0317] ：
2008-10-02 13 : 11:06 发表#
莎朗斯 的 B 都 给 别人 操 烂 了 ！ **1**
网易 四川 资阳 网友 ip ： 222.213 . * . * ：
2008-08-27 23 : 28:26 发表#
地震 的 那 一个 星期 我 根本 就 不敢 看 电视 ， 因为 看 一次 就要 哭 一次 ， 不想 第二天 上班 的 时候 红 红肿 肿 的 。 没 想到 听到 那个 老 女人 这样 称 四川 地震 有趣 ！ 天 啦 ！ 这 都 是 人 说 的 说 吗 ？ 本来 迪奥 这个 品牌 的 东东 我 一直 都 在用 的 ， 现在 开始 ！ ！ ！ ！ 从今 以 后 ！ ！ ！ ！ ！ ！ ！ 坚决 抵制 法国 迪奥 化妆品 ， 莎朗斯通 的 影片 ！！ **2**
网易 重庆 永川 网友 ip ： 222.181 . * . * ：
2008-06-10 12 : 11:02 发表#
网易 加拿大 网友 (70.69 . * . *) 的 原 贴 ：
大家 团结 起来 ， 让 她 知道 中国 人民 的 力量 ！ 坚决 抵制 法国 迪奥 化妆品 (莎朗斯通 代言) ， 莎朗斯通 的 影片 ！！
网易 辽宁 大连 网友 (116.3 . * . *) 的 原 贴 ：

Sur le forum, les intervenants s'identifient par un numéro IP[12]. Les propos sont parfois exprimés avec certaine retenue, parfois avec une grande violence, n'excluant pas la vulgarité :

1. 莎朗斯 的 B 都 给 别人 操 烂 了 !

   Sharon Stone a été b… jusqu'à la destruction de son s… !

2. 本来 迪奥 这个 品牌 的 东东 我 一直 都 在用 的， 现在 开始！ ！ ！ ！ 从今 以后！ ！ ！ ！ ！ ！ 坚决 抵制 法国 迪奥 化妆品， 莎朗斯通 的 影片!!

   Jusqu'à présent, j'utilisais assez souvent les produits Dior, à partir de maintenant !!!! Et dorénavant !!!!!!! Je les boycotterai totalement ainsi que les films de Sharon Stone !!

3.1 Segmentation du texte

Les comparaisons textométriques supposent que l'on définisse des unités de décompte dont on étudie ensuite les variations de fréquence au sein des différentes parties du corpus. Pour pourvoir mettre en oeuvre des comparaisons textométriques, on utilise des outils informatiques qui permettent de découper automatiquement les unités du texte avec lesquelles on pourra s'en servir de leurs occurrences et spécificités.

[12] Notons que ce numéro d'identification, avant tout destiné à l'identification de la machine utilisée sur l'Internet pour des satisfaire des impératifs techniques de transfert de données, peut également être utilisé pour localiser l'internaute sur le web avec une précision plus ou moins élevée.

莎朗•斯通放厥词竟称地震是报应 引发广泛批评

2008年05月27日 15:52:46 来源：北京晨报

戛纳电影节昨天凌晨落幕，却有一丝不和谐声音传出。美国女星莎朗•斯通日前接受采访时出言不逊，引发多方批评。其代言的法国奢侈品品牌迪奥（Dior）昨天紧急表态，称迪奥公司绝对不认同莎朗•斯通的言论，并将严肃考虑这一事件，稍后会发布公开声明。

5月22日，莎朗•斯通出席戛纳为艾滋病筹款的慈善晚会，在红地毯上被记者问起对四川汶川地震的看法。莎朗•斯通先是称自己“不喜欢中国”，地震很“有趣”，随后又称被地震中的人和事感动，考虑为地震灾区“做点什么”。

虽然莎朗•斯通最后试图表达对灾区的同情，但是她言语中带出的幸灾乐祸和不屑一顾的意味引起了大众的强烈反感。这段视频被传到网上后，一天之内点

莎朗 斯 通 放 厥 词 竟 称 地震 是 报应 引发 广泛 批评
2008年 05月 27日 15: 52: 46 来源 ： 北京 晨报

戛 纳 电影节 昨天 凌晨 落幕 ， 却 有 一 丝 不 和谐 声音 传出 。 美国 女星 莎朗•斯通 日前 接受 采访 时 出言不逊 ， 引发 多方 批评 。 其 代言 的 法国 奢侈品 品牌 迪奥 （ Dior ） 昨天 紧急 表态 ， 称 迪奥 公司 绝对 不 认同 莎朗•斯 通 的 言论 ， 并 将 严肃 考虑 这 一 事件 ， 稍后 会 发布 公开 声明 。

5月 22日 ， 莎朗•斯 通 出席 戛纳 为 艾滋病 筹款 的 慈善 晚会 ， 在 红 地毯 上 被 记者 问 起 对 四川 汶川 地震 的 看法 。 莎朗•斯 通 先是 称 自己 “ 不 喜欢 中国 ”， 地震 很 “ 有趣 ”， 随后 又 称 被 地震 中 的 人 和 事 感动 ， 考虑 为 地震 灾区 “ 做 点 什 么 ”。

虽然 莎朗•斯 通 最后 试图 表达 对 灾区 的 同情 ， 但是 她 言语 中 带 出 的 幸灾乐祸 和 不屑一顾 的 意味 引起 了 大众 的 强烈 反 感 。 这 段 视频 被 传 到 网上 后 ， 一 天 之内 点 击 率 达到 20 多 万 次 ， 留言 超过 1600 条 ， 大多数 留言 都 批评 莎朗•斯 通 的 说法 。 网友 hairychest23 留言 称 ： “ 按照 她 的 逻辑 ， 当年 发 生 在 美国 新 奥尔良 的 飓风 也 不 是 自然灾害 ， 而是 某种 报应 了 ？ ”网友 willshine 说 ： “ 斯通 女士 ， 你 居然 可以 看 着 无

Traduction : Sharon Stone a mis en relation le tremblement de terre avec un mauvais Karma : Ceci a déclenché de nombreuses critiques 27 mai 2008 15:52:46 Source: Beijing Morning News

Le festival de Cannes a pris fin hier matin dans le calme. Quelques jours avant, l'actrice américaine Sharon Stone avait fait des déclarations brutales dans une interview. Ceci a déclenché de multiples critiques. La marque de luxe *Dior* dont elle est la représentante a déclaré hier que la société Dior se désolidarisait des propos tenus par Sharon Stone, et qu'elle envisageait de donner une suite sérieuse à cet incident. Dior fera une déclaration publique ultérieurement.

Le 22 mai à Cannes, Sharon Stone a participé à un gala de charité pour la lutte contre le sida. Sur le tapis rouge elle a été interviewée par un journaliste à propos du tremblement de terre à Sichuan. Tout d'abord, Sharon Stone a déclaré qu'elle « n'aime pas trop la Chine », le tremblement de terre était « intéressant », puis elle a dit qu'elle est touchée par les gens et les événements du tremblement de terre, également, elle envisage « faire quelque chose » pour la zone frappée par le séisme. /…/

Figure 5

Exemple d'un fragment de presse extrait du journal *Beijing Morning News* avant (à gauche) et après segmentation en mots par le logiciel ***Haylanda*** (à droite) suivi de la traduction en français du début de l'extrait (en bas du tableau).

Si la notion de *mot* est bien définie dans les grammaires chinoises, l'écriture chinoise n'intègre pas d'espace entre les unités lexicales. Les lecteurs chinois appréhendent les textes en découpant la chaîne textuelle en unités distinctes, à partir de leurs propres connaissances linguistiques. Les mots chinois sont composés d'un à quatre sinogrammes. Un même caractère peut avoir différentes fonctions grammaticales en fonction de son contexte.

Cette particularité constitue une difficulté spécifique pour l'exploitation textométrique des textes chinois. Pour pouvoir découper les textes en unités correspondant plus ou moins à des mots du chinois, nous avons choisi le logiciel de segmentation ***Haylanda*** [13]. On trouvera, ci-dessous les principales caractéristiques lexicométriques des trois sous-ensembles rassemblés dans le corpus ***StoneKarma***.

Tableau 2

Principales caractéristiques lexicométriques
des trois sous-ensembles du corpus ***StoneKarma***

Partie	occurrences	formes	hapax	F. Max
Blog	21538	3987	2161	1703
Forum	158132	7277	3209	7815
Presse	17937	2643	1055	1188

On se gardera d'interpréter directement des différences entre ces caractéristiques textométriques qui peuvent résulter d'artéfacts produits par des méthodes d'encodage et de stockage différents pour chacun des supports rassemblés en un même corpus.

3.2 Comparaisons entre médias

Les trois volets diffèrent, bien entendu par les types d'expression propres à chaque média. Style plutôt soutenu et tournures officielles pour la presse, style plus littéraire mais aussi plus personnel pour l'expression sur les blogs, très grande variété de modes d'expression sur les forums qui peuvent aller d'un style soutenu à des listes d'interjections et d'insultes.

Statut de la répétition dans les trois volets

Les procédures de repérages des segments répétés (suite de formes reproduites à l'identique à différents endroits du corpus) permettent de repérer des reprises de séquences plus ou moins étendues dans chacun des trois volets du corpus.

Sur le tableau 5, on peut voir la répétition de séquences localisées au sein d'articles de presse. Cette répétition vient ici de la reprise fréquente, par les différents journaux, du même texte présentant, au nom de Sharon Stone, des excuses exprimées par son agent :

> */…/ Mes propos déplacés ont blessé le peuple chinois et ont suscité sa colère. Je m'en excuse profondément./…/*

Les commentaires qui accompagnent ce texte contiennent aussi des répétitions de phrases de commentaires et de réactions souvent identiques.

[13] Ce logiciel est en développé par la société Hailanda Segmentation intelligente, version d'essai (海量智能分词研究版 http://www.hylanda.com/).

Tableau 3 :
Exemple de répétition localisée dans deux articles de presse

<p=P12> # 莎朗 - 斯通 道歉 遭 中国 网友 拒绝 外 媒 称 其 是 白痴 <ANNEE="2008"><MOIS="2008_5"><JOUR="2008_5_29"> 04 : 35 据 上海 《 青年报 》 报道， “ 我 的 错误 言行， 让 中国 人民 感到 悲伤 与 愤怒， 我 为此 深感 歉意 。” 这 是 莎朗 - 斯通 于 昨日 下午 通过 其 经纪人 发给 Dior 公司 的 致歉 文， 对于 自己 于 5 月 24 日 在 法国 戛纳电影节 上 发表 的 关于 中国 四川 地震 的 冷血 言论 作出 忏悔， 并 表示 愿意 积极 参与 救援 活动 。 。 。 。 。 。 。 。 。 。 。 。 。 。 。 。 昨晚， 莎朗 - 斯 通通 过 经纪人 向 其 代言 的 品牌 公司 发 了 一封 致歉 信， 称 由于 自己 在 受访 中 的 不当 言行， 对 中国 人民 所 造成 的 伤害， 感到 非常 的 难过 和 ??歉 。 同时， 莎朗 - 斯 通称， 对于 此次 自然灾害 中 所有 遭受 不幸 的 受害者， 表示 最 深切 的 ??悼 与 同情 。 “ 我 的 错误 言行， 让 中国 人民 感到 悲伤 与 愤怒， 我 为此 深感 歉意 。 我 再次 强调， 我 愿意 积极 参与 任何 关于 中国 地震 灾害 的 援助 活动 , 并 尽全力 帮助 受灾 的 中国 人民 。” 虽然 道歉 信 的 内容 颇 显 其 忏悔 之 意， 然而， 截至 记者 截稿 时， 莎朗 - 斯通 本人 及 其 经纪 公司 并未 有 任何 要 为 中国 地震 捐款 或 捐物 等 所谓 的 救援 行为， 道歉 目前 还 只是 停留 在 书面 形式 。
<p=P12> # Les excuses de Sharon refusées par les internautes chinois. Les médias étrangers la qualifient d « imbécile ». <ANNEE="2008"><MOIS="2008_5"><JOUR="2008_5_29"> 4 : 35. Selon les reportages du « Journal des jeunes » de Shanghaï, « *Mes propos déplacés ont blessé le peuple chinois et ont suscité sa colère. Je m'en excuse profondément* » (reproduit 2-3 fois dans différents articles) /.../

On peut voir au tableau 4 une intervention due à un internaute qui manifeste son indignation en reproduisant un grand nombre de fois la même séquence d'injures : *Sharon Stone, tu es une espèce de chienne ; crève* ! Comme on le comprend aisément, ce type de répétition est facilité par les possibilités technologiques de l'écriture électronique (l'internaute a visiblement construit son message à l'aide du *copier/coller* répété plusieurs fois) qui rend ces répétitions incomparablement plus faciles à produire que lors d'un énoncé vocal, par exemple.

Tableau 4 :
Exemple de répétition réalisée sur un forum de discussion à base de copier/coller

<ANNEE="2008"><MOIS="2008_05"><JOUR="2008_05_28">22 : 12 : 17 发表# 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 莎朗斯通 你 个 老 母狗 去 死 吧 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 莎朗斯通 你 个 老 母狗 去 死 吧 网易 美国 网友
<ANNEE="2008"><MOIS="2008_05"><JOUR="2008_05_28">> l'IP n° 22 : 12 : 17. x .x dit : Sharon Stone, tu es une espèce de chienne ; crève (répété 25 fois)

Le statut de ces répétitions dans les comptages textométriques se complique d'autant. D'un côté, il n'est pas tout à fait naturel de considérer que ces nombreuses répétitions équivalent à une seule séquence (l'internaute a dupliqué son texte 25 fois, ce qui traduit très certainement une grande irritation de sa part), d'un autre côté, la prise en compte de ces 25 occurrences

d'une même séquence ne peut être mise sur le même plan que 25 occurrences d'une même séquence produites par des individus différents (avec ou sans citation mutuelle).

4 Etude contextuelle de la forme 抵制-(di zhi boycott)

L'étude des *spécificités maximales* (mots particulièrement sur-employés) pour le volet ***forum*** met en évidence un emploi massif du terme : 抵制, (di zhi) par les internautes qui s'expriment sur ce type de support. Le terme 抵制 (di *résister*+ zhi *maîtrise*) correspond plus ou moins au terme occidental de *boycott*. Il apparaît avec une fréquence très élevée (784 occurrences) dans ce volet du corpus, immédiatement après les particules grammaticales qui sont les mots les plus fréquents du corpus.

Nous avons tenté d'analyser l'emploi de cette forme dans le corpus ***StoneKarma*** sous le triple aspect de sa répartition à l'intérieur des textes, de la comparaison de ses contextes immédiats et de la liste des formes lexicales qu'il attire dans son entourage syntagmatique.

4.1 Répartition de la forme

L'histogramme de la figure 6 montre la répartition du terme 抵制 (di zhi, boycott) au sein des trois ensembles de textes rassemblés dans le corpus. Comme nous l'avons signalé plus haut, le terme apparaît très majoritairement dans la partie *forum* du corpus (733 occurrences dans les forums sur 784 au total).

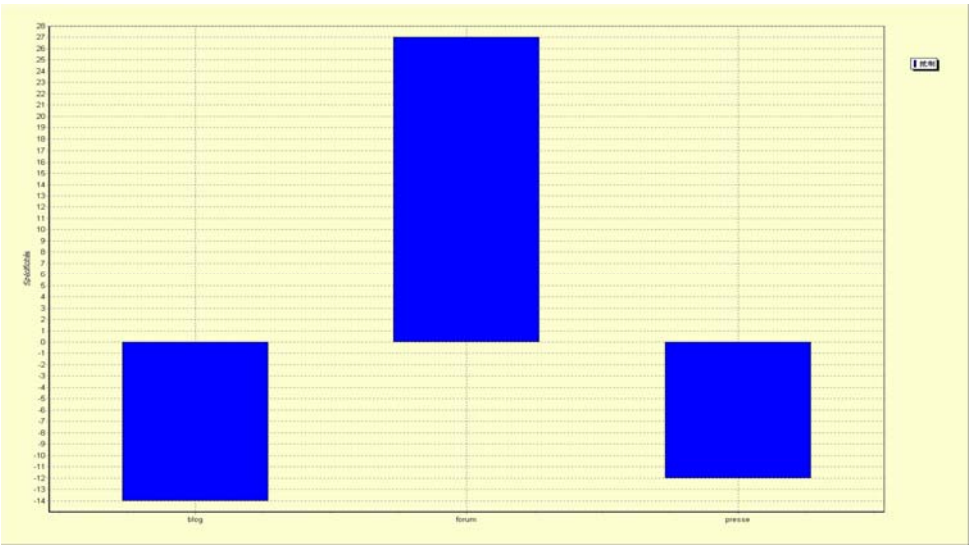

Figure 6 :
Ventilation de la forme 抵制 (boycott)
dans les 3 parties du corpus ***StoneKarma***

Cet emploi privilégié n'est pas sans rapport avec les mécanismes de saturation des messages à l'aide du *copier/coller* dont nous avons présenté un exemple ci-dessus (un même paragraphe pouvant contenir un nombre important des occurrences du terme). Cependant, la disproportion en faveur des forums nous amène à conclure que ce mot trouve une faveur particulière chez les intervenants des forums, alors que les rédacteurs de presse et de blogs, sans doute tenus à une certaine réserve évitent de l'employer trop souvent.

4.2 Contextes

La figure 6 montre, pour chaque média sélectionné, un certain nombre de contextes dans lesquels on retrouve la forme 抵制 (di zhi *boycott*). Comme on le voit, les contextes de cette forme repris dans les articles de presse proviennent le plus souvent de discours rapportés dont les auteurs sont des citoyens que l'on interroge dans le cadre du reportage :

> , 并 呼吁 全国 所有 书店 、 音像店 共同 抵制 莎朗 - 斯通 。 记者 昨日 采访 的 重庆 某
> Appelons au boycott des produits de Sharon Stone dans toutes les librairies et boutiques
> , 引得 华人 世界 震怒 。 网友 一致 呼吁 抵制 莎朗 斯通 代言 产品 。 昨天 下午 , 其
> C'est un grand choc pour le monde sinophone, les internautes appellent au boycott de Sharon Stone ainsi que des produits qu'elle représente

Les contextes prélevés sur des forums résultent au contraire de l'expression directe d'un appel au boycott, dont la cible peut varier, de la part des citoyens chinois.

> 果 DIOR 不 更换 代言人 , 大家 起来 坚决 抵制 DIOR, 大家 行动 起来 网易 湖北 黄石 网友
> Unissons-nous pour boycotter DIOR s'il ne change pas de représentant. 我们
> 必须 抵制 DIOR , 只有 这样 , 才能 让 她 付出 代价
> C'est du mépris pour la vie humaine, nous devons boycotter DIOR ! Elle doit le payer !
> 斯通代言 的 任何 产品 ! ! 山东 人民 坚决 抵制 SBST[14] 的 任何 电影 , 包括 其 代言 的 任何
> Boycott de tous les films et les produits de cette conne de Sharon Stone

L'inventaire distributionnel réalisé après la même forme permet de hiérarchiser les entités que les internautes proposent de soumettre à un boycott.

Inventaire distributionnel	*Equivalent français*
415 ---- ---- --- 抵制 法国	boycott France
2 ---- ---- --- 抵制 和	boycott et
2 ---- ---- --- 抵制 美国	boycott USA
2 ---- ---- --- 抵制 你	boycott toi
8 ---- ---- --- 抵制 其	boycott ce qu'elle
7 ---- --- 抵制 其 代言	boycott de tout ce qu'elle représente
6 --- 抵制 其 代言 产品	boycott les produits qu'elle représente
10 ---- ---- --- 抵制 莎朗	boycott Sharon
2 ---- ---- --- 抵制 莎朗·斯通	boycott Sharon Stone
40 ---- ---- --- 抵制 莎朗斯通	boycott Sharon Stone
9 ---- --- 抵制 莎朗斯通 代言	boycott de ce qu'elle représente
4 --- 抵制 莎朗斯通 的 影片	boycott les films de ST
2 --- 抵制 莎朗斯通 所有	boycott de tout ce que
2 ---- ---- --- 抵制 他	boycott (elle)

Tableau 5 :
Extrait de l'inventaire distributionnel après la forme 抵制 (di zhi, *boycott*)

[14] SBST, initiales en caractères romains de *Shabi 傻屄 Sharone Stone, conne de Sharon Stone.*

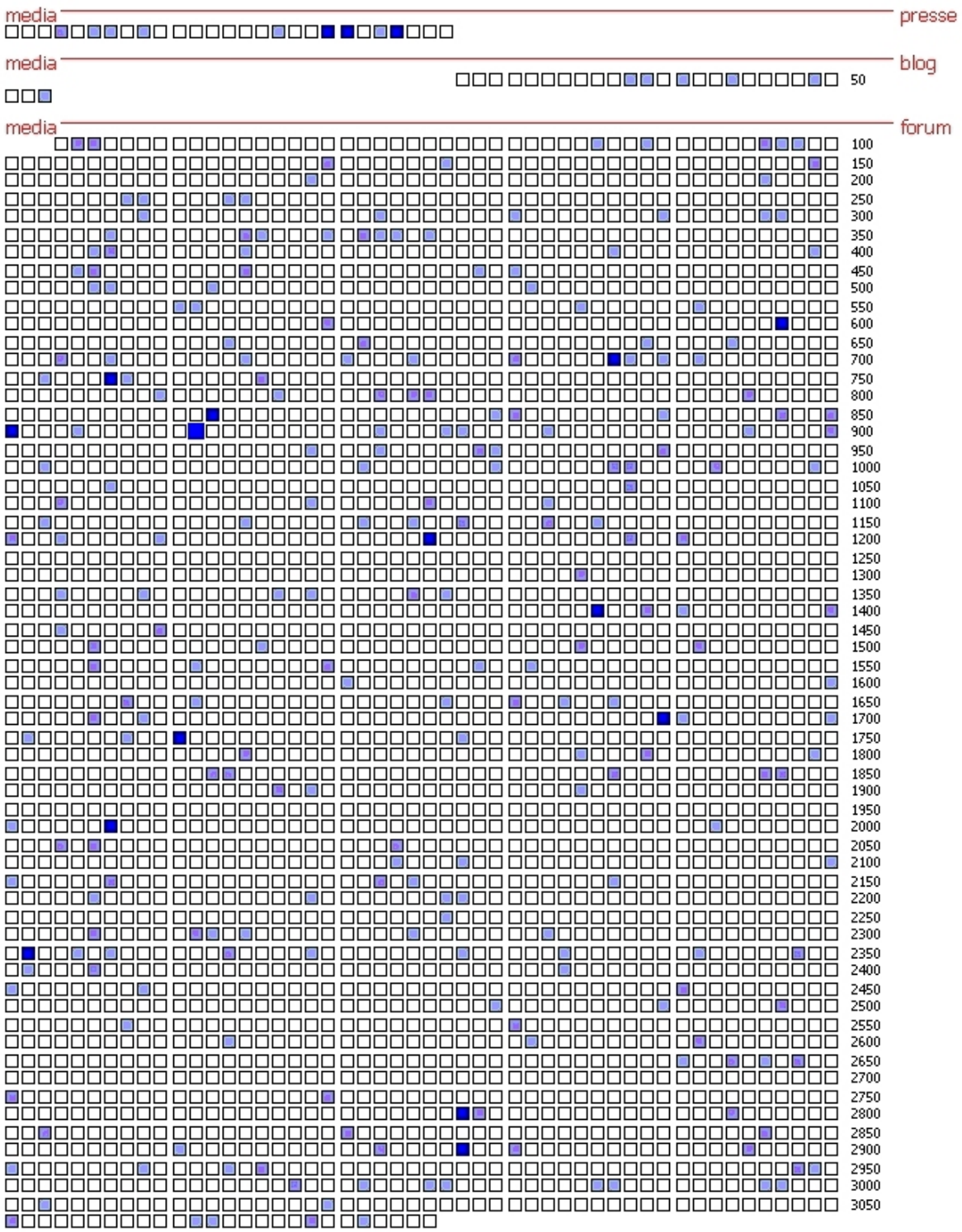

Figure 7 :

Carte des sections réalisées à partir des occurrences
de la forme 抵制 (di zhi, *boycott*)

La carte des sections qui montre la distribution de cette même unité à l'intérieur des sections (paragraphes) découpées dans le corpus permet de localiser cette vision avec une plus grande précision et de vérifier que le terme, outre ses emplois massifs par certains des internautes

repérables par la couleur foncée des paragraphes qui correspondent à leurs interventions, est largement utilisé par un grand nombre d'intervenants.

Guide de lecture pour la figure 7

Dans la carte des sections qui correspond au corpus StoneKarma, les volets correspondant à chacun des médias étudiés sont séparés par une ligne rouge.
Pour chacun des volets, les différentes sources (presse, blogs, forums) sont représentées par un carré.
Les carrés de couleur vive permettent de repérer les sections qui utilisent particulièrement le mot pour lequel la carte a été établie (ici la forme 抵制 di zhi, *boycott*).

Contextes de 抵制 (di zhi, boycott) dans la partie PRESSE (tri après)

有 网友 提出 拒 看 ‘ 莎 剧 抵制 其 代言 产品 ， 得到 了 大多数 人 的 ??
? 愤怒 ， 并 呼吁 中国 人群 起 发动 全面 抵制 莎朗 - 斯通 。 香港 / 文汇报 / 报道
， 并 呼吁 全国 所有 书店 、 音像店 共同 抵制 莎朗 - 斯通 。 记者 昨日 采访 的 重庆 某
品 重庆 下 架 据 记者 了解 ， 为了 抵制 莎朗 - 斯通 不 负 责任 的 讲话 ， 近日 已
， 引得 华人 世界 震怒 。 网友 一致 呼吁 抵制 莎朗 斯通 代言 产品 。 昨天 下午 ， 其
丛?： 网易 娱乐 ） 上海 书店 抵制 莎朗 - 斯通 影视 作品 全部 下 架 2008 年
UME 院线 的 老板 吴思远 就 公开 表示 要 抵制 莎朗 - 斯通 主演 的 电影 。 莎朗 - 斯通
UME 院线 的 老板 吴思远 就 公开 表示要 抵制 莎朗 - 斯通 主演 的 电影 。 莎朗 - 斯通

Contextes de 抵制(di zhi , boycott)dans la partie FORUM (tri après)

果 DIOR 不 更换 代言人 ， 大家 起来 坚决 抵制 DIOR ， 大家 行动 起来 网易 湖北 黄石 网友
?， 是 对 人类 生命 的 漠视 . 我们 必须 抵制 DIOR ， 只有 这样 ， 才能 让 她 付出 代价
- 27 16 : 21 : 11 发表 # 抵制 ， 坚决 抵制 ， SB 货 网易 上海 黄浦 网友 ip : 58 . 37
. * . * ） 的 原 贴 ： 河北 人民坚决 抵制 SBST 的 任何 电影 ， 包括 其 代言 的 任何
?代言 的 任何 产品 ！ ！ 河北 人民 坚决 抵制 SBST 的 任何 电影 ， 包括 其 代言 的 任何
?代言 的 任何 产品 ！ ！ 山东 人民 坚决 抵制 SBST 的 任何 电影 ， 包括 其 代言 的 任何
190 . * . * ） 的 原 贴 ： 辽宁人民坚决 抵制 SBST 的 任何 电影 ， 包括 其 代言 的 任何
泄?人民 的 力量 ！ ！ ！ 辽宁 人民 坚决 抵制 SBST 的 任何 电影 ， 包括 其 代言 的 任何
* ： 2008 - 05 - 27 17 : 24 : 3 发表 抵制 SBST 的 任何 电影 网易 江苏 泰州 网友 ip
?出来 ， 公布 到 各个 论坛 。 我们 共同 抵制 。 北京 、 上海 的 书店 今天 已经 把 ??
5 - 27 21 : 31 : 45 为 有 默默 的 抵制 。 网易 江西 新余 网友 ip : 218 . 64 .
表 # 他妈的 ， 私通 ， 政治 娼妇 坚决 抵制 。 网易 上海 杨浦 网友 [wang13386190]
2008 - 05 - 27 21 : 46 : 50 坚决 抵制 “ 私通 ” 影片 ， 在 网上 查出 她 的 影片

Contextes de 抵制(di zhi, boycott)dans la partie BLOGS (tri après)

法国 奢侈品 品牌 迪奥 （ DIOR ） 已 被 抵制 。 绝大部分 的 民众 意见 都 是 希望 立刻
立即 引发 了 国人 对 莎朗 - 斯通的全面 抵制 。 同样 “ 有趣 ” 的 是 ， 莎朗 - 斯通
锟?表示 ： “ 她 的 音像制品 也 应 遭到 抵制 。” 演艺圈 愤怒 声讨 ， 要 其 道歉
代言 的 DIOR 产品 ， 昨天 不少 网友发 抵制 ， 以 表 愤怒 。 对此 ， 记者 第一 时间
要 把 这个 产品 列入 抵制 的 黑名单， 抵制 到 该 品牌 撤销 与 其 合同 为止 ? ? 这样
营销 ， 我们 也 要 把 这个 产品 列入 抵制 的 黑名单 ， 抵制 到 该 品牌 撤销 与 其
. 全面 封杀 莎朗斯通 的 电影 。 5 坚决 抵制 购买 莎朗斯通 代言 的 一切 周边 产品 ！
?莎朗斯通 ， 华人 娱乐圈 也 已经 开始 抵制 莎朗斯通 影视 作品 ， 香港 节目 知名 主持人
， 国内 数 以 十万 网民 们 联合发起了 抵制 沙朗斯通 的 大 反击 ！ 而 还 在 法国 戛
维护 人类 的 基本 良知 。我 从来不提倡 抵制 什么 国家 的 品牌 和 产品 。 但 此时此刻
拇?， 我们 中国 人民 不 欢迎 你 ， 坚决 抵制 它 代言 的 所以 产品 ， 不许 它 的 产品
国 必须 封杀 她， 有 良知 的 人 都应该 抵制 她。 刘威 ： 这样 的 艺人 根本 不

Figure 6 :

Extraits de la concordance de la forme 抵制 (di zhi *boycott*)
dans chacune des trois parties du corpus ***StoneKarma***

On voit sur l'inventaire distributionnel réalisé après la forme 抵制 (boycott/boycotter) que l'objet de l'action de boycott envisagée est prioritairement la France ou les produits français (415 occurrences). Les USA n'apparaissent qu'occasionnellement dans ce contexte. La personne de Sharon Stone elle-même n'est visée que secondairement.

Cooccurrences

La recherche des cooccurrences (i.e. formes de vocabulaire apparaissant plus souvent qu'un modèle de répartition homogène ne le laisserait supposer) à l'intérieur des seuls textes produits sur les forums permet de préciser encore l'objet du boycott prôné par les intervenants sur les forums.

Parmi les segments les plus longs et les plus répétés dans le volet forum du corpus, on repère aisément des séquences en forme de mot d'ordre qui permettent de mieux cerner la nature de la colère exprimée par les internautes.

Forme	Equivalent français	Tot.	Fréq.
的 力量 ！ 坚决 抵制	*boycotter fermement*	379	379
坚决 抵制 法国 迪奥 化妆品	*boycotter les cosmétiques de Dior France*	413	413
大家 团结 起来,让 她 知道 中国 人民 的 力量	*unissons-nous ; pour faire connaître la force du peuple chinois*	377	377

6 Conclusion

La détection ou la fouille d'opinions est un domaine de recherche en plein essor. Ils peuvent se révéler cruciales pour les entreprises et trouve de très nombreux domaines d'applications veille technologique, marketing, concurentielle, études politiques et sociétales. La mise en relation des opinions et sentiments exprimés avec les thèmes sur lesquels ces opinions et sentiments portent est encore un domaine en cours d'exploration, dont les enjeux concernent la transformation des informations extraites des textes en informations structurées en connaissances synthétisées et exploitables.

À partir d'une vive polémique déclanchée dans l'opinion publique chinoise par les propos d'une actrice américaine, nous nous sommes intéréssés à la partie du débat accessible sur le réseau internet.

Utilisant les deux grands moteurs de recherche (Google et Baidu) nous référencer, via un certain nombre de mots-clés les textes les plus consultés par les internautes et relatifs à ce débat.

Dans un second temps, nous avons sélectionné, pour trois types de supports médiatiques identifiés (presse, blogs, forums) des échantillons de texte qui nous ont permis de constituer un corpus d'étude. Ce type d'étude, surtout lorsqu'il est pratiqué sur un échantillon restreint, ne saurait se présenter comme une synthèse des réactions repérables dans l'opinion publique chinoise.

Les traitements textométriques nous ont permis de constater la diversité des réactions exprimées par les internautes et de découvrir une hiérarchie inattendue des thèmes exprimés grâce aux observations sur la diversité des réactions.

Nous pensons avoir montré la possibilité qu'il y a d'accéder à des réactions authentiques vis-à-vis de ce qu'on croit de la liberté d'expressions, exprimées dans trois différentes couches de l'opinion publique chinoise.

7 Références

Lamalle C., Salem A., "Types généralisés et topographie textuelle dans l'analyse quantitative des corpus textuels", Actes *des 6èmes Journées d'analyse des données textuelles,* St Malo, 2002

Lebart L., Salem A., *Statistique textuelle*, Paris, Dunod, 1994, téléchargeable sur le site : http://www.cavi.univ-paris3.fr/lexicometrica/livre/st94/st94-tdm.html

Miao J., Salem A., Comparaisons textométriques de traductions franco-chinoises, in *Explorations textométriques*, 2008.

Shen L., http://lionelshen.free.fr/Labo/Master/Memoire_M2_LS.pdf