

## **Le référencement de sites multilingues (2ème partie : Japon et Chine)**

[Retour au sommaire de la lettre](#)

*De nombreux sites sont ciblés vers plusieurs populations, parlant des langues différentes. Le mois dernier, nous tentions d'apporter un éclairage précis sur les différentes façons de structurer un site web en fonction de plusieurs zones linguistiques, depuis la meilleure solution, jusqu'à... la pire ! ce mois-ci, nous explorons deux pays très dynamiques sur le Web : le Japon et la Chine. Comment référencer son site sur les moteurs de recherche leaders dans ces deux pays ? Voici quelques indications et voies à suivre...*

Le référencement multilingue pourrait presque devenir un pléonasme à l'avenir, tant les langues diverses sont présentes aujourd'hui dans les échanges professionnels. Tout d'abord, parce que l'entreprise évolue dans une Europe à 27 (et presque autant de langues !), ensuite parce qu'Internet favorise largement les échanges commerciaux entre les 27 pays européens et bien entendu au delà de ses frontières, entre continents.

C'est sur ce dernier point que nous avons souhaité nous pencher plus particulièrement sur quelques spécificités d'un référencement à destination du continent asiatique. Nous nous intéresserons, dans cet article, à deux pays à l'économie très dynamique, notamment sur le Web : la Chine et le Japon.

### **Quelles contraintes pour les langues asiatiques ?**

Dans ce contexte, l'enjeu d'une présence en première page des moteurs de recherche est essentiel mais demande au référencement un travail d'adaptation aux "contraintes" de ces pays. Contraintes d'ordre :

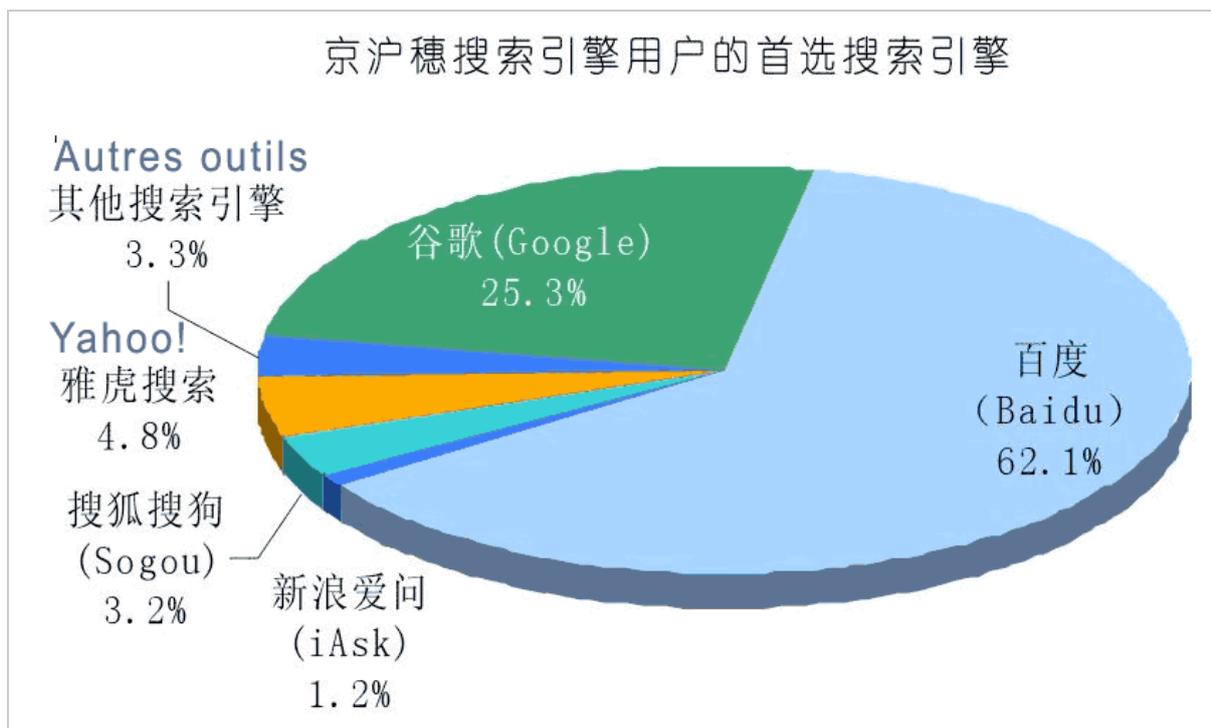
- Commercial : quel est le paysage des outils de recherche dans le pays ciblé, les offres disponibles ?
- Technique : les processus d'indexation des pages d'un site et de calcul de pertinence en vue de leur positionnement comportent-ils des changements par rapport à l'Europe ?
- Sémantique : comment construire les contenus ? Quelles expressions choisir ? Comment ventiler ces expressions dans ces contenus ?

### **Choix du nom de domaine**

Dans la première partie du dossier consacré au référencement multilingue, nous avons vu toute l'importance de bien choisir la solution technique la plus favorable au bon positionnement du site, chacune présentant ses forces et ses faiblesses. Nous considérerons ici la solution la plus "profitable", à savoir un nom de domaine par langue. En effet, pour rappel, cette "indépendance" présente, entre autres, deux avantages déterminants : une adaptation culturelle graphique du site, très importante lorsque l'on s'adresse à une population aux repères différents, mais également un meilleur traitement des pages du site par les moteurs, notamment lors des recherches locales (exemple : recherche de pages en chinois en provenance de Chine, voir première partie de l'article le mois dernier).

### **Paysage des outils de recherche en Chine**

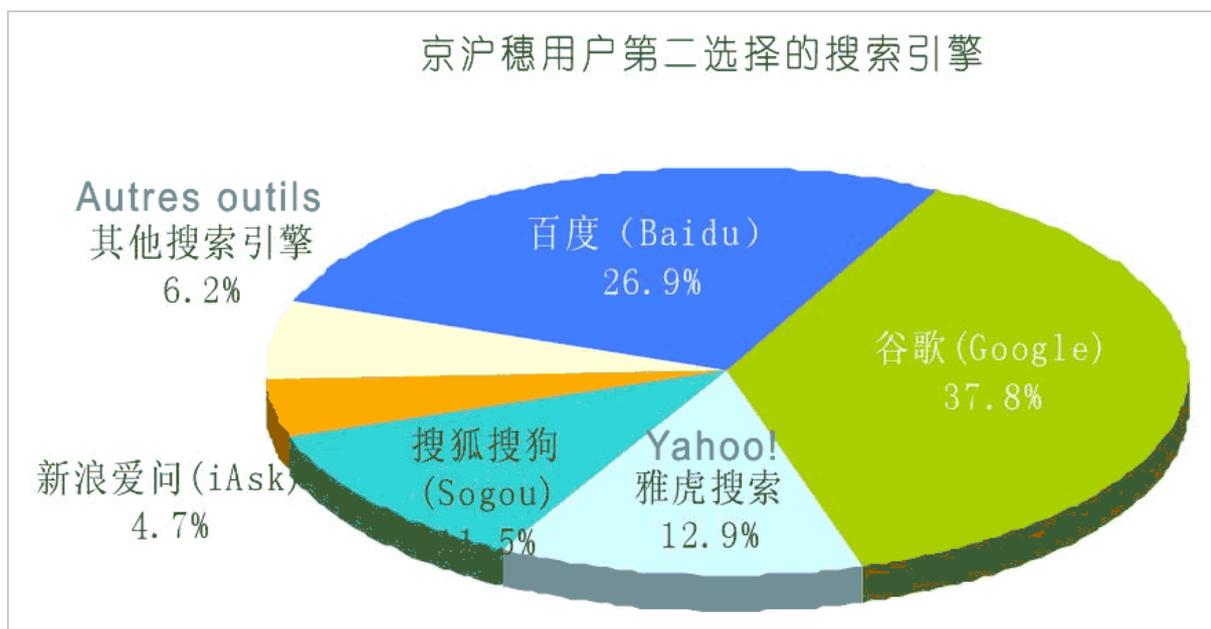
Le premier choix de moteur de recherche par les internautes chinois est indiqué sur le graphique ci-dessous (concernant des internautes de Pékin, Shanghai, Guangzhou, source : <http://www.cnnic.cn/>, Septembre 2006, China Internet Network Information Center) :



*Premier choix de moteur des internautes chinois*

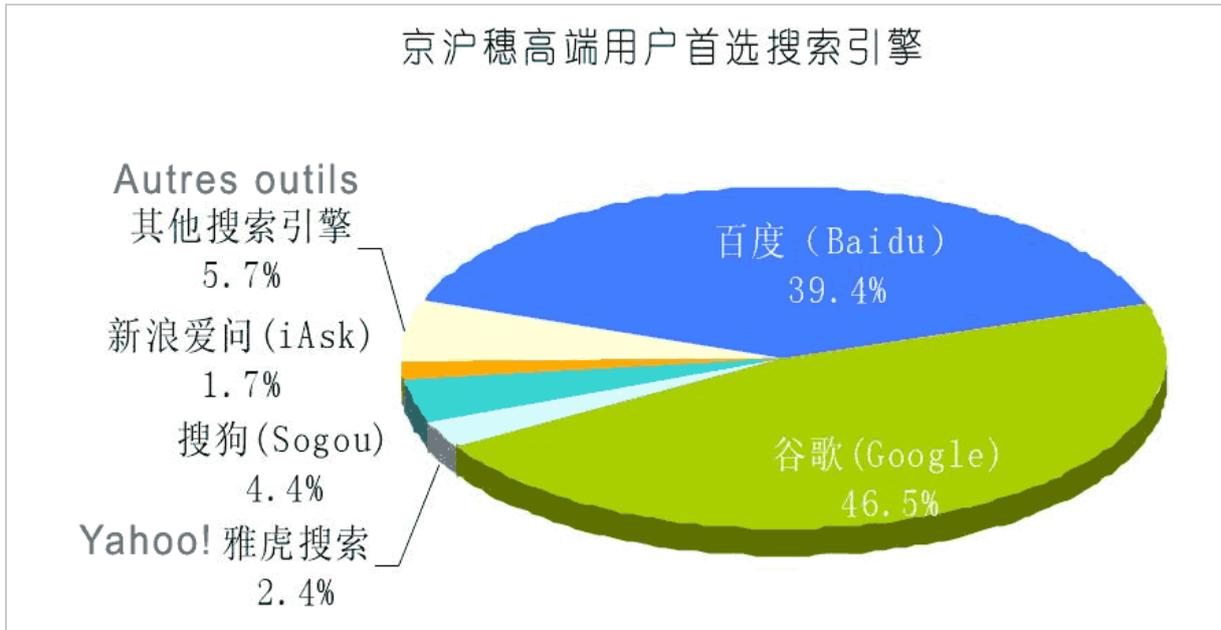
Contrairement à l'Europe où Google dispose d'une position hégémonique, le marché des moteurs est beaucoup plus disputé en Chine, notamment avec des moteurs locaux comme Baidu (<http://www.baidu.com/>), qui occupe la position de leader.

76% des internautes chinois utilisent deux moteurs de recherche. Leur deuxième choix se porte alors, logiquement, sur Google :



*Deuxième choix de moteur des internautes chinois*

Le choix de l'"élite" se porte, lui, sur Google. Le sondage définit l'"élite" des internautes chinois comme suit : plus de 25 ans, bac + 4, revenu mensuel supérieur à 3000 Yuans, non-étudiant... L'"élite" représente 19,3% des internautes non-étudiants, et 11,2% du nombre total d'internautes.



Choix de moteur par l'"élite" des internautes chinois

**Baidu : priorité au référencement payant !**

Comme Google, Baidu propose, à droite des résultats de recherche, des liens sponsorisés, mais leur positionnement est régi par un abonnement annuel dans le moteur. Les prix d'abonnement sont définis par Baidu en fonction de la concurrence et du positionnement sur un mot-clé pour une période d'un an. Les trois premières positions sont toujours fixes. Les positions suivantes en revanche varient selon un roulement déterminé par Baidu.

La différence entre Baidu et ses homologues américains se situe au niveau de la partie gauche de la page de résultats. En effet, tout comme le moteur Voila.fr en France, les premiers résultats présentés à l'internaute sont issus des liens commerciaux, dont la position est fonction du prix de l'enchère. Ces liens commerciaux sont facturés au prix au clic (PPC).

Cela se passe d'ailleurs à l'insu de la majorité des internautes chinois, puisqu'aucune mention ne stipule le caractère payant de ces résultats !

Exemple : pour la requête "tête de taraudage" : **攻丝机**

**Liens commerciaux : Positionnement via un abonnement annuel**

**Liens sponsorisés : Positionnement au coût par clic**

21 entreprises participaient à cette enchère au mois de janvier 2007, le prix au CPC variant de 2,85 à 0,30 Yuan. Cela signifie que le référencement organique est exclu des deux premières pages ! Une donnée qu'il sera important de prendre en compte au moment de définir votre stratégie de référencement dans ce pays et plus spécifiquement sur cet outil !

### Principales différences de fonctionnement des outils de recherche chinois

Nous ne disposons pas d'informations officielles précises sur le fonctionnement de l'indexation organique par les robots de Baidu. Toutefois, il est expliqué dans leurs pages, au même titre que Google d'ailleurs, que le "spam" doit être banni, que le contenu doit être pertinent, que le nombre de caractères dans les titres (balises TITLE) est limité, et que l'on peut s'exposer à la sanction de la "Sandbox", etc.

Concernant Google Chine, il semble que le fait de suivre les règles éditées dans les "Guidelines" du site Google occidental (<http://www.google.com/support/webmasters/>) reste une stratégie gagnante.

Par exemple pour la requête "Europe Web marketing" - "欧洲网络营销" Activis est positionné en premier sur 3 400 000 résultats, sur la base d'une optimisation proche de ce qui peut se faire en Occident :

The screenshot shows a Google search interface in Chinese. The search bar contains the text "欧洲网络营销". Below the search bar, there are radio buttons for "所有网页", "中文网页", "简体中文网页", and "中国的网页". The search results are displayed in a list format, with each result including a title, a snippet, and a URL. The first result is from "www.activis.net/simplifiedchinese/search-engine-marketing-europe.htm" with a title "欧洲搜索引擎营销, 欧洲搜索引擎关键词排名竞价, 关键词广告, 提升网站...". The second result is from "www.avenit.com.cn/page\_438\_cn.html" with a title "蔚尼公司- 欧洲电子商务解决方案, 全方位的欧洲互联网服务, 互联网服务...". The third result is from "www.avenit.com.cn/page\_432\_cn.html" with a title "蔚尼公司- 欧洲电子商务解决方案, 全方位的欧洲互联网服务, 互联网服务...". The fourth result is from "www.activis.net/simplifiedchinese/news\_qualified\_google.htm" with a title "Activis, 欧洲网络营销科技权威: 建设多国语言优质网站激活全球网络营销".

Il est particulièrement difficile de décrypter la manière dont Google établit la relation entre les mots d'une expression. Tout comme pour l'encodage des caractères chinois. Pour l'heure, nous n'avons pas trouvé d'éléments concluant de différenciation avec le fonctionnement du Google occidental. Une optimisation "asiatique" s'effectuera donc de la même façon, globalement, qu'une optimisation "occidentale"...

### Différence au niveau du traitement sémantique

Chinois simplifié vs. chinois traditionnel : si vous connaissez un peu l'histoire et la culture chinoise, l'existence de deux versions de la langue ne vous est pas étrangère. Cela date de la révolution

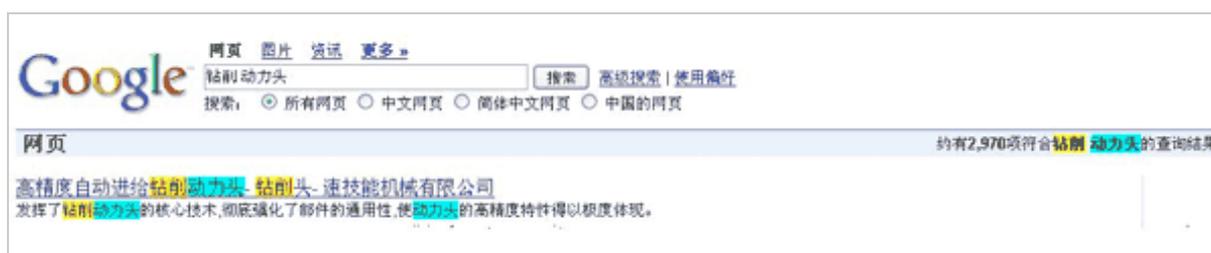
culturelle imposée dans les années 60, mais ce n'est pas le sujet ici. Toujours est-il qu'aujourd'hui, les deux versions sont utilisées et doivent cohabiter (le traditionnel est utilisé à Hong Kong, à Macao et à Taiwan et le simplifié est utilisé en République Populaire de Chine et à Singapour).

L'intérêt pour le référencement est donc de créer deux versions du site, l'une en chinois simplifié (汉字简体中文), l'autre en chinois traditionnel (漢字繁體中文), accessibles par un lien hypertexte simple ou un drapeau. Les deux versions seront donc indexées séparément. L'avantage est conséquent : il s'agit de gagner une double visibilité auprès des internautes chinois.

### Question de sémantique ?

Bien que Baidu soit le premier moteur de recherche en Chine, Google ne perd pas la confiance des internautes chinois. Il faudra donc prendre en compte, en majorité, ces deux outils. Or, outre la question de la censure (qui, là aussi, est un autre sujet., non moins important.), on constate des différences de logique des analyses sémantiques.

Prenons un exemple dans le domaine de l'outillage : "unité de perçage", qui peut s'écrire "钻削动力头" en chinois simplifié et "动力钻销头" en traditionnel.



Expression "unité de perçage" en chinois simplifié ("钻削动力头") : résultat de recherche sur Google Chine : 198 000, sur Baidu 2 550.



Expression "unité de perçage" en chinois traditionnel ("动力钻销头") : résultat de recherche sur Google Chine : 197 000, sur Baidu 6.

Apparemment les combinaisons d'expressions ne sont pas traitées avec la même logique (vraisemblablement au niveau des opérateurs booléens "or", "and", etc.) par les deux moteurs de recherche. Autre explication possible : les deux géants diffèrent au niveau de la taille de leur index et la différence semble aujourd'hui très importante.

## Référencement japonais

Tout comme en Chine, Google ne domine pas le marché japonais de la recherche. En revanche, dans ce pays, c'est Yahoo qui mène la danse et non un moteur local.

		PPC deta
Yahoo!Japan	64.8%	GoogleAdwords/Overture
Google	17%	GoogleAdwords
MSN search	9.4%	Overture
goo	2.1%	GoogleAdwords
Biglobe	2%	GoogleAdwords
excite	1.8%	GoogleAdwords
@nifty	1%	GoogleAdwords
infoseek	1%	GoogleAdwords
AOL Search	0.4%	GoogleAdwords
freshEYE	0.2%	Overture
So-net	0.1%	GoogleAdwords

Source : <http://www.japan-seo.cn/japanse-market.htm>

D'autres chiffres, plus récents mais moins précis (les statistiques Alexa sont tirées de l'utilisation de sa barre d'outils, sans que l'on sache si celle-ci est réellement très prisée au Japon) :

[http://www.alexa.com/site/ds/top\\_sites?cc=JP&ts\\_mode=country&lang=none](http://www.alexa.com/site/ds/top_sites?cc=JP&ts_mode=country&lang=none)

## Principales différences de fonctionnement des outils de recherche japonais

Encodage html des caractères japonais : bien entendu, lorsque vous mettez en ligne une version de votre site en japonais, il est nécessaire que le moteur puisse suivre les liens vers cette version, à l'aide d'expressions-clés. L'encodage des caractères japonais répond à certaines spécificités et peut se faire de différentes manières : en UTF-8, en EUC-JP et en Shift\_jis. Or, chaque moteur peut se limiter à un certain type d'encodage pour l'affichage des caractères japonais. Ainsi Google qui aime les standards semble préférer l'UTF-8, alors que Yahoo, lui, est plus sensible au Shift\_jis et à l'EUC-JP.

Optimisation des noms de domaines et des urls : latin quand tu nous tiens ! Comme en chinois, les noms de domaine et urls sont aujourd'hui uniquement disponibles en alphabet latin. N'espérez donc pas optimiser vos urls en japonais, cela ne fonctionne pas encore, même si la porte commence à s'ouvrir. Lire à ce sujet :

- Vers une internationalisation des noms de domaines sur le web - Actualités - ZDNet.fr

<http://www.zdnet.fr/actualites/internet/0,39020774,39364550,00.htm>

- Des caractères chinois dans l'adresse

<http://pisani.blog.lemonde.fr/2006/12/08/des-caracteres-chinois-dans-ladresse/>

Cela étant, vos concurrents étant dans la même situation, cela n'est pas trop pénalisant...

Concernant, les noms de domaine, s'ils ne sont pas optimisables, ils peuvent cependant influencer la qualité du référencement. Ainsi, il semble que si Google ne se soucie pas trop de l'extension du nom de domaine pour bien référencer une page en japonais, Yahoo, quant à lui, paraît donner la préférence aux site en .jp.

## Traitement sémantique : les spécificités "littéraires" du japonais :

Particularités grammaticales : du côté littéraire, là encore c'est le latin qui parfois nous joue des tours et ce, aussi bien sur Google que sur Yahoo.

Ainsi une expression comme "mesure de niveau" ne s'écrira pas en trois mots, mais en un seul tel que "mesureniveau" (soit レベル計 en japonais), "sonde de mesure de niveau capacitif" s'écrira "capacitifniveauètre" (soit 静電容量式レベル計 en japonais).

Or, en terme de référencement un mot unique ne sera pas correctement identifié s'il est "collé" à un autre. Ainsi dans 静電容量式レベル計 (sonde de mesure de niveau capacitif), l'expression "mesure de niveau" (レベル計) n'aura que très peu de poids pour les moteurs de recherche.

Ajoutez à cela les trois modes d'écritures japonaises qui se mélangent (Katakana, Hiragana, et Kanji), on se trouve finalement dans un type de référencement où le seul choix des mots clés selon leur "sens européen" ne suffit pas. Exemple :

- "mesure de niveau" s'écrit : レベル計

- "niveau" s'écrit en Katakana : レベル

- "mesure" en Kanji : 計, mais ce terme peut aussi désigner l'outil lui-même qui sert à faire la mesure, etc...

Il faut donc en plus, savoir comment les mots-clés sont écrits, combien il existe de manières de les écrire et enfin si cela est compatible avec les exigences du référencement (et on ne parle pas ici de la ponctuation...).

Vous l'aurez compris, le référencement en langue asiatique est passionnant !

Pour finir sur une note humoristique, voici une vidéo qui présente le fonctionnement pour le moins surprenant de l'outil de recherche japonais GOO. Excellent !

<http://www.youtube.com/watch?v=cqen8H1PNzw>

**Article proposé par David Eichholtzer, Judy Chen Ferlay et Simon Vissol – Chargés de référencement au sein de la société Activis – <http://www.activis.net/>  
Activis est spécialiste du référencement multilingue  
[http://www.activis.net/dn\\_resultats\\_referencement\\_multilingue/](http://www.activis.net/dn_resultats_referencement_multilingue/) et traite plus de 12 langues dont les langues asiatiques :  
[http://www.activis.net/dn\\_referencement\\_asiatique/](http://www.activis.net/dn_referencement_asiatique/)**