

**Identification des leaders d'opinion sur internet :
utilisation des données secondaires issues de Twitter**

JEAN-PHILIPPE GALAN

Maître de Conférences

Centre de Recherche en Management, EAC 5032 CNRS

IAE de Toulouse – Département Marketing

Université de Toulouse Capitole

2, rue du Doyen Gabriel Marty

31042 Toulouse Cedex 9

e-mail : jean-philippe.galan@iae-toulouse.fr

ALEXANDRA VIGNOLLES

Enseignant-Chercheur en Marketing

Insec Business School

26, rue Raze

33000 Bordeaux

e-mail : avignolles@groupeinsec.com

Identification des leaders d'opinion sur internet : utilisation des données secondaires issues de Twitter

Le développement de l'internet et de manière générale des technologies de l'information et de la communication permet plus que jamais aux utilisateurs de produire du contenu et d'entrer en interaction. Les forums de discussion, les blogs, les sites de réseaux sociaux sont des services – regroupés sous le terme de Web 2.0 – participant à cette nouvelle utilisation collaborative du média internet. Parmi les conséquences de cette augmentation des échanges d'information par voie électronique se trouve le regain d'intérêt pour le bouche à oreille (Larceneux, 2007) et le leadership d'opinion (Vernette, 2006). Plus que jamais, le consommateur peut communiquer avec les autres, bien au-delà des limites des relations qu'il pouvait entretenir jusque là avec son entourage, créant ainsi une « agora à l'échelle mondiale » (Florès, 2008 ; p.8).

Dès les années 50, Katz et Lazarsfeld (1955) ont suggéré que les échanges avec l'entourage étaient plus persuasifs que les communications en provenance des entreprises via les médias de masse. Ainsi les entreprises auraient tout intérêt à communiquer auprès de certains individus appelés leaders d'opinion qui propageraient ensuite le message commercial à leur entourage de manière plus efficace (Holmes et Lett, 1977). Cette approche se retrouve sur l'internet où l'on parle d' « e-fluentials » (Cakim, 2006) ou d'« influenceurs » (Vernette et Florès, 2004) pour désigner ces utilisateurs possédant une influence sur leur réseau relationnel *on-line*. Il est dès lors stratégique pour l'entreprise de repérer ces individus qui participent entre autres à la diffusion des innovations ou à l'anticipation des préférences des consommateurs (Vernette, 1997).

L'objectif de cette communication est de proposer l'utilisation des données secondaires issues du service Twitter pour repérer les influenceurs sur l'internet. Après une présentation rapide de l'intérêt de cette méthode par rapport aux méthodes traditionnelles d'identification, nous détaillerons les spécificités de ce service et présenterons les résultats d'une étude conduite sur 2000 individus. Cette étude débouche sur la proposition d'un indice unique pour repérer les influenceurs dont le principal avantage est sa gratuité et sa simplicité d'accès et de calcul.

1 – Cadre théorique : Leadership d’opinion et internet.

Il existe plusieurs méthodes d’identification des leaders d’opinion dans un contexte traditionnel. Celles-ci font régulièrement l’objet d’évaluation (e.g. Vernet et Giannelloni, 2004 ; Bertrandias, 2004). Quatre techniques sont régulièrement citées : (1) L’observation qui consiste à examiner l’activité d’un groupe ou d’une communauté afin d’analyser les relations entre les membres et la dynamique des échanges. (2) La sociométrie qui consiste à demander aux membres d’une communauté de désigner les leaders. (3) L’informateur clé qui consiste à interroger des membres particulièrement informés sur un groupe d’individus et leur demander de décrire les relations des membres. (4) L’auto-désignation, méthode la plus utilisée en marketing, dans laquelle les individus évaluent eux-mêmes l’influence qu’ils exercent sur leur entourage.

Si l’influence interpersonnelle sur internet est un domaine qui semble bénéficier d’un fort engouement (e.g. Vernet et Florès, 2004 ; Cakim, 2006 ; Belvaux et Marteaux, 2007 ; Larceneux, 2007, Gouldenberg et al., 2009), il n’en est pas de même des études ayant pour objectif de proposer des méthodes pour repérer les influenceurs. A l’heure actuelle la seule véritable tentative est celle de Fejlaoui et Vernet (2009) qui s’appuie sur des méthodes netnographiques (Kozinets, 2002) et d’auto-désignation. Leur travail débouche sur la proposition d’une échelle de mesure du leadership des membres de communautés *on-line*. Si la méthode d’auto-désignation est précise et riche d’implications, elle pose cependant des contraintes de réactivité, de coût et d’accès aux données dont une approche exclusivement fondée sur les données secondaires permet de se libérer. En effet, la construction d’une échelle de mesure nécessite du temps (processus de génération d’items, épuration, validation ou *a minima* une adaptation au contexte si l’échelle préexiste à l’étude) ce qui parfois peut freiner le décideur nécessitant des résultats à court terme. C’est une procédure coûteuse demandant des budgets pour la collecte, le traitement et l’analyse. Enfin, elle dépend de la bonne volonté des individus à accepter de répondre à l’étude ce qui génère d’emblée un biais dans l’évaluation de la portée des résultats. Quoique sur l’internet, il est apparu dans un premier temps que ces contraintes pouvaient paraître moins fortes (Galan et Vernet, 2000), la forte sollicitation des individus à participer à des enquêtes par ce média a fortement contribué à gommer cet avantage (Florès, 2008).

Aujourd'hui, l'internet, grâce à l'avènement du Web 2.0, regorge de données qui sont produites en réaction des comportements des consommateurs (Galan, 2002 ; Roxin et Bernard, 2007 ; Florès, 2008) et dont l'accès est la plupart du temps gratuit. Le principal avantage est qu'elles sont acquises par l'observation et non par le questionnement ce qui présente l'avantage de ne pas poser de problème de recrutement de participants et diminue les biais. Certains ont d'ailleurs développé des méthodologies très pertinentes de ce type de données à l'instar de la netnographie (Kozinets, 2002). Cependant, dans beaucoup de cas, si l'accès aux données est simplifié, le traitement et l'analyse de données demeurent un travail long et coûteux, et cela même si certains logiciels peuvent permettre d'alléger la charge de travail (Gauzente, 2006).

Pour nous, la démocratisation du micro-blogging, *via* le service Twitter qui a vu sa notoriété et le nombre de ses utilisateurs fortement progresser durant ces derniers mois, permet de fournir un nouveau type de données secondaires qui permet d'alléger encore les contraintes qui pèsent sur les études *on-line* (Galan et Vignolles, 2009). Dans la section qui suit, nous allons exposer rapidement les spécificités de ce service avant de présenter une application de l'analyse de données secondaires sur Twitter dans le but de repérer les influenceurs.

2 – Méthodologie : Les données secondaires fournies par Twitter

Apparu en 2006, Twitter a récemment connu une forte progression de notoriété et d'utilisateurs. L'IFOP rapporte que la notoriété de ce service a été multipliée par 7 durant le premier semestre 2009 pour arriver à plus de 9 millions de français ayant entendu parler de Twitter et estime le nombre d'utilisateurs abonnées à 640 000 (Cann, 2009). Ce service gratuit de micro-blogging et de réseau social propose à ses utilisateurs de diffuser du contenu textuel sous la contrainte que celui-ci ne dépasse pas 140 caractères (d'où le terme de micro-blogging). Ces courts messages sont dénommés « tweets » et peuvent être diffusés via un ordinateur ou un téléphone mobile. A l'instar du blog, il permet donc à ses utilisateurs de s'exprimer auprès de leur « communauté ». Le positionnement initial de Twitter était de proposer un service permettant à ses utilisateurs de tenir en permanence leur entourage de leurs faits et gestes. L'utilisation a cependant considérablement évolué et permet à tout individu de diffuser des informations sur n'importe quel sujet, donner son avis sur l'actualité ou encore de confesser des idées ou réflexions.

Twitter est également un service de réseau social : pour lire les « tweets » d'un individu il faut le suivre, c'est-à-dire s'inscrire sur le réseau Twitter et s'abonner au flux que propose l'utilisateur. Contrairement au réseau Facebook qui nécessite un accord de l'auteur des contenus (statuts Facebook), sur Twitter, l'utilisateur n'a pas besoin de l'accord de l'émetteur des « tweets » pour les lire. Par défaut, chaque utilisateur accepte que quiconque le suive, c'est-à-dire ait accès au contenu qu'il produit. Le service donne toutefois l'opportunité de paramétrer le compte de manière à ce qu'une demande soit acceptée pour le suivi, mais très peu d'utilisateurs ont recours à cette option. A posteriori, un utilisateur peut bloquer un autre utilisateur, l'empêchant ainsi de le suivre. Enfin, l'utilisateur suivi reçoit un courrier électronique le notifiant que quelqu'un le suit. Il peut alors choisir de le suivre en retour, se retrouvant alors dans une relation mutuelle de suiveur-suivi¹ et symétrique comparable à celles qui se nouent sur le réseau Facebook.

De ce fait, il est aisé de suivre plusieurs utilisateurs que l'on ne connaît pas. De la même manière la procédure pour arrêter de les suivre se fait en un « clic » et ne nécessite aucun accord. Ceci a pour conséquence que, même si pour chaque utilisateur le contenu est archivé tout comme dans un blog, l'accès aux informations se fait en temps réel auprès de tous les utilisateurs suivis en même temps plutôt que par la consultation de la page « personnelle » de chacun. Le service se présente ainsi comme une page actualisée en permanence sur laquelle défilent tous les « tweets » produits par les individus que l'utilisateur a décidé de suivre. Cette particularité de focaliser l'utilisateur sur l'instant présent a rapidement fait de Twitter un concurrent redoutable pour tous les services d'actualités. Twitter possède en effet un fort avantage vis-à-vis des sites traditionnels d'information : l'information est livrée en temps réel par les utilisateurs, de manière brute. Alors que d'autres doivent faire des choix éditoriaux pour choisir quelles informations méritent plus que les autres d'être communiquées aux lecteurs, Twitter les fournit toutes. Cet apparent fatras a poussé les utilisateurs à s'adapter afin de faciliter l'accès à cette masse d'information. Sont alors apparus les hashtags et les retweets.

Les hashtags sont des termes qui, à l'intérieur des tweets, sont précédés d'un symbole dièse (#) et associent le tweet à un concept plus général, à une catégorie, ce qui permet de faciliter le traitement de l'information (Ladwein, 1995). Ainsi la présence des hashtags #politique #cuisine etc permet immédiatement de contextualiser le message et d'en faciliter le traitement.

¹ Twitter n'étant pas encore disponible en langue française, les termes utilisés sont « follower » et « following » que nous traduirons donc dans cette communication par suiveur et suivi.

Tout comme les tags constitutifs des folksonomies (Roxin et Bernard, 2007), ces hashtags peuvent être créés sur simple décision de l'utilisateur sans faire référence à une taxonomie officielle des sujets abordés sur Twitter.

Les retweets (RT) sont une information qui permet à l'utilisateur de citer sa source lorsqu'il reprend à son compte l'information d'un autre utilisateur de Twitter. Il précise alors qu'il n'est pas la première source de l'information en indiquant « RT @ » suivi du pseudonyme de l'utilisateur source : « RT @toto : le texte hadopi2 a finalement été voté ». Ce faisant, le RT marque l'intérêt porté par un utilisateur au contenu d'un autre utilisateur et s'impose au lecteur comme le signe de la valeur informationnelle particulière du message.

Ainsi, dans le flux d'informations qui défile sur son écran, l'utilisateur peut rapidement et visuellement repérer les thématiques (#) et les messages importants (RT). Cette immédiateté de l'information brute, partiellement catégorisée et hiérarchisée par les retweets et les hashtags combinée à la fonction de réseau social du service Twitter paraît particulièrement intéressante pour étudier les phénomènes de bouche à oreille et d'influence interpersonnelle sur l'internet. En effet, le fait qu'il n'y a pas une nécessité d'accord entre utilisateurs pour se suivre, mutuellement ou non, permet à certaines informations de circuler de très nombreuses fois, véhiculées d'un utilisateur à l'autre. Un message peut-être à l'origine produit par un individu ayant peu d'utilisateurs abonnés à son flux. Pourtant si l'un d'eux choisit de « retweeter » une de ses informations, il peut se créer une propagation très forte amplifiée par le nombre de suiveurs de chaque utilisateur ayant « retweeté » l'information. On comprend bien alors que certaines personnes ont un vrai poids dans la circulation de l'information et cela explique comment l'annonce de la mort de Michael Jackson avait fait en quelques minutes le tour de la « twittosphère » bien avant la première dépêche officielle.

L'objectif de ce travail étant de proposer une méthode pour repérer les influenceurs sur l'internet, la section suivante va s'attacher à présenter les éléments qui permettent de construire un indicateur de l'influence des utilisateurs de Twitter et enfin de vérifier son pouvoir discriminant dans la population étudiée.

3 – Résultats

Nous avons suivi deux mille personnes durant neuf semaines. Seules des personnes déclarant résider en France et s'exprimant en français ont été sélectionnées. Sur les 2000 personnes suivies, 53 (2.65%) nous ont bloqué, c'est-à-dire empêché de les suivre une fois qu'elles ont pris connaissance de notre suivi et 9,15% des utilisateurs suivis nous ont suivi en retour. Ceci nous amène à 183 relations symétriques soit 9,4%. Enfin, la simple création de notre compte pour les besoins de cette étude et le fait qu'en tant qu'utilisateur nous figurions dans des listes de suiveurs ou de suivis d'autres utilisateurs a amené certains à nous suivre spontanément (97 personnes). La collecte a été assurée par un programme de récupération des données sur l'API (Application Programming Interface de Twitter) créé spécialement pour les besoins de cette étude.

Nous rappellerons que Twitter est un service jeune qui a connu une forte notoriété en France principalement ces derniers mois. La population étudiée reflète cet état de fait. En effet, beaucoup d'utilisateurs sont nouveaux sur Twitter. Le tableau 1 détaille l'ancienneté des membres de notre échantillon. Le constat est que sur ces 1947 personnes, au début de l'enquête, la moitié utilisaient le service Twitter depuis moins de trois mois.

Tableau 1 – Ancienneté des individus de l'échantillon

| Ancienneté | Fréquence | % | Cumulé | |
|--------------|-----------|--------|-----------|---------|
| | | | Fréquence | % |
| < 1 semaine | 65 | 3,34% | 65 | 3,34% |
| < 2 semaines | 148 | 7,60% | 213 | 10,94% |
| < 3 semaines | 156 | 8,01% | 369 | 18,95% |
| < 4 semaines | 123 | 6,32% | 492 | 25,27% |
| < 3 mois | 422 | 21,7% | 914 | 46,94% |
| < 6 mois | 357 | 18,3% | 1271 | 65,28% |
| < 1 an | 294 | 15,10% | 1565 | 80,38% |
| <3 ans | 382 | 19,6% | 1947 | 100,00% |

3.1 – Repérer les influenceurs

L'objectif de cette étude est de proposer à la fois une méthodologie pour repérer les influenceurs sur Twitter et de vérifier que les personnes ainsi détectées présentent des

caractéristiques cohérentes avec ce statut présumé. La première question est donc : comment repérer les influenceurs parmi les 1947 (soit 2000 – 53) utilisateurs suivis ?

Pour qu'un individu soit un influenceur, il faut qu'il ait quelqu'un à influencer. Le nombre de suiveurs paraît donc être un candidat potentiel au titre d'indicateur du leadership. On remarque effectivement (cf. tableau 2) que beaucoup d'individus sont très peu suivis (près d'un tiers de l'échantillon est suivi par moins de 15 personnes et près de la moitié est suivi par moins de 50 personnes) alors qu'un très petit nombre d'individus est suivi par un très grand nombre de personnes (une cinquantaine d'individus est suivi par plus de 1000 suiveurs) allant jusqu'à plus de 9000 dans notre échantillon.

Tableau 2 – Nombre de suiveurs et de suivis par individu et en % de la population

| Nombre de suiveurs par individu | | | Nombre de suivis par individu | | |
|---------------------------------|-----------|---------|-------------------------------|-----------|---------|
| suiveurs | individus | % | suivis | individus | % |
| 10 | 370 | 19,00% | 10 | 253 | 12,99% |
| 15 | 542 | 27,84% | 20 | 488 | 25,06% |
| 25 | 750 | 38,52% | 30 | 643 | 33,03% |
| 50 | 1075 | 55,21% | 40 | 774 | 39,75% |
| 100 | 1370 | 70,36% | 50 | 894 | 45,92% |
| 300 | 1740 | 89,37% | 150 | 1454 | 74,68% |
| 500 | 1831 | 94,04% | 200 | 1567 | 80,48% |
| 1000 | 1907 | 97,95% | 300 | 1694 | 87,01% |
| 2000 | 1934 | 99,33% | 500 | 1807 | 92,81% |
| 3000 | 1937 | 99,49% | 1 000 | 1885 | 96,82% |
| 4000 | 1940 | 99,64% | 2 000 | 1936 | 99,44% |
| 9500 | 1947 | 100,00% | 10 000 | 1947 | 100,00% |

Le nombre de suiveurs apparaît *a priori* comme un premier indicateur de l'influence d'un individu sur Twitter. Il est même un critère souvent cité par les professionnels pour signifier l'influence d'un individu (Vuitton, 2009). Pourtant, un rapide coup d'œil sur la structure des données montre que la structure de suivi est très similaire. En effet, le tableau 2 montre bien qu'à l'instar du nombre de suiveurs, un grand nombre d'individus suivent assez peu de personnes (25% des individus suivent moins de 20 personnes, 75% en suivent moins de 150)

alors qu'un petit nombre d'individus suit un grand nombre de personnes (moins de 5% de l'échantillon suivent plus de mille personnes).

Comme nous l'avons expliqué plus haut, une certaine proportion d'individus suivent en retour qu'ils sont suivis, créant ainsi une relation entre le nombre de suivis et le nombre de suiveurs. En outre, plus l'individu suit d'autres personnes et plus son nom figure de nombreuses fois sur Twitter (du fait qu'il apparaît dans la liste de chaque personne suivie). Etant donné que le principal moyen de suivre d'autres utilisateurs de Twitter est de cliquer sur leur nom dans les listes de suiveurs ou de suivis, ceci augmente encore l'opportunité de faire augmenter le nombre de suiveurs.

Il ressort du tableau 2 que le nombre de personnes suivies a un impact direct sur le nombre de suiveurs. Nos données confirment une corrélation de 0,796 ($p < ,000$) entre les deux variables. Ceci corrobore le fait que plus l'individu suit de gens et plus il est suivi en retour, ce qui ne traduit pas forcément un leadership particulier mais plutôt une mécanique due au fonctionnement de Twitter. Le nombre de suiveurs n'apparaît finalement pas pertinent comme indicateur du leadership de l'utilisateur. Aussi, nous pensons que des individus qui présentent un grand nombre de suiveurs et de suivis ne sont pas réellement à considérer comme des influenceurs. Ceux qui ont un grand nombre de suiveurs et de suivis sont ceux qui ont le plus de liens sociaux sur Twitter, ils seront donc qualifiés de « connecteurs », terme proposé par Goldenberg *et al.* (2009) pour signifier des « personnes ayant un nombre exceptionnellement grand de liens sociaux ».

3.2 – Le ratio suiveurs/suivis (S/s) comme indicateur du leadership

S'il existe une très forte corrélation entre le nombre de suiveurs et le nombre de suivis, il n'en reste pas moins que certains individus sortent clairement de ces corrélations et on peut observer certains individus « atypiques » qui ont un rapport suiveurs / suivis très éloigné de 1. D'un côté certains individus suivent beaucoup de monde et sont très peu suivis en retour. Même si cette communication n'est pas le lieu pour discuter ce profil, nos données montrent qu'il s'agit principalement de « jeunes comptes » qui suivent en attendant d'être suivis, et de personnes n'offrant aucun contenu (pas ou peu de « tweets »), bloquant le suivi en retour, typiques donc *a priori* de personnes utilisant Twitter comme source d'information plutôt que pour son potentiel relationnel.

D'un autre côté certaines personnes affichent un rapport suiveurs/suivis fortement supérieur à 1 pouvant aller jusqu'à 105 dans nos données. Il y a également deux types de comptes correspondant à ce profil. D'une part ce que l'on nomme des célébrités ou « people » (acteurs, chanteurs, artistes, hommes politiques, présentateur tv, membre d'une émission de tv réalité, etc) qui utilisent Twitter pour donner des actualités sur leur vie personnelle ou le calendrier de leurs déplacements et/ou productions. D'autre part, des individus qui ne jouissent pas d'une popularité particulière en dehors de Twitter ou du réseau internet (blog, site web, etc). C'est à cette dernière partie de notre échantillon que nous nous sommes intéressés.

Ce qui vient intuitivement à l'esprit, lorsque l'on se demande comment certains peuvent avoir beaucoup plus de suiveurs que de suivis, c'est que la production de contenu (nombre de tweets et qualité) doit en être à l'origine (e.g. Gilly *et alii*, 1998). Sur Twitter il apparaît clairement qu'un nombre restreint d'individus est à l'origine de la majorité du flux d'informations. Ceci pourrait donc confirmer que la production d'information pourrait être à l'origine du leadership sur l'internet comme l'a déjà suggéré Kozinets (2002) dans un autre contexte (newsgroups).

Tableau 3 : nombre de tweets par individu

| individus | % | tweets | % |
|------------------|----------|---------------|----------|
| 18 | 1 | 26 488 | 17,62 |
| 45 | 2,5 | 45 001 | 29,94 |
| 90 | 5 | 63 358 | 42,16 |
| 180 | 10 | 85 970 | 57,2 |
| 360 | 20 | 111 355 | 74,23 |
| 540 | 30 | 126 503 | 84,17 |
| 720 | 40 | 136 125 | 90,57 |
| 900 | 50 | 142 067 | 94,53 |
| 1350 | 75 | 148 980 | 99,12 |
| 1799 | 100 | 150 295 | 100 |

Il est important de noter que 148 (7,55%) comptes sont totalement inactifs. Parfois anciens (plus d'un an), il ne s'y passe absolument plus rien. Si l'on ne tient pas compte de ces utilisateurs qui ont peut-être ouvert un autre compte, perdu leur identifiant, la vie ou quoique

ce soit d'autre les empêchant d'utiliser Twitter, la population passe à 1799 utilisateurs actifs. Sur les deux mois d'observation, 150 295 « tweets » ont été produits par les utilisateurs. Le nombre de « tweets » par individu est récapitulé dans le tableau 3. S'il apparaît effectivement que certains produisent plus de contenu que d'autres, ceci ne semble pourtant pas avoir un effet déterminant sur le ratio suiveurs / suivis. Le tableau 4 nous aide à mieux comprendre les liens entre variables.

On remarque que si la corrélation est significative du fait du grand nombre d'observations, celle-ci est très basse pour être proche de zéro. Le nombre de tweets est en revanche fortement corrélé au nombre de suiveurs ($r=,374$) et de suivis ($r=,288$). Il semble donc que la production de contenu soit une cause de l'augmentation du nombre de suiveurs et probablement une conséquence de l'augmentation du nombre de suivis. Quoiqu'il soit difficile de dire exactement la nature profonde de ces relations, il paraît évident que plus l'individu va être enchevêtré dans un réseau relationnel et plus il va produire de contenu en réponse aux contenus qu'il va recevoir de son réseau.

Tableau 4 : corrélation du ratio S/s avec les suiveurs, suivis, tweets et ancienneté

| | Ratio S/s | Nb suiveurs | Nb suivis | Nb tweets | Ancienneté |
|-------------|-----------|-------------|-----------|-----------|------------|
| Ratio S/s | 1 | ,415* | -,26 | ,066* | ,068* |
| Nb suiveurs | ,415* | 1 | ,796* | ,374* | ,188* |
| Nb suivis | -,026 | ,796* | 1 | ,288* | ,121* |
| Nb tweets | ,066* | ,374* | ,288* | 1 | ,016 |
| Ancienneté | ,068* | ,188* | ,121* | ,016 | 1 |

* $p < ,001$

Il est évident voire mécanique que le nombre de suiveurs soit corrélé au ratio suiveurs / suivis. Aussi, une fois enlevé cette variable, on remarque que le ratio suiveurs/suivis est difficilement explicable par les autres éléments constitutifs du compte et une régression multiple montre que si l'ancienneté ($t=2,108$; $p < ,035$) et le statut ($t=1,947$; $p < ,047$) ont un effet significatif sur le ratio suiveurs / suivis ($F(2 ; 1796) = 6,144$; $p < ,002$), le pouvoir explicatif de cette relation est très faible ($R^2 = ,007$).

Le ratio suiveurs/suivis qui montre le rapport d'influence que l'individu entretient avec son réseau relationnel n'est pas explicable mécaniquement par le fonctionnement du réseau Twitter. Il semble donc qu'il traduise une véritable caractéristique de l'individu. A cet instant de nos investigations, il semble être un indicateur très pertinent pour le repérage des influenceurs.

3.3 – Pertinence du ratio suiveurs/suivis

Alors qu'il peut être très compliqué de mesurer l'impact de l'influence des leaders d'opinion dans d'autres contextes, il émerge de la communauté des utilisateurs de Twitter une fonctionnalité qui permet d'apprécier l'influence des individus : le retweet. Le retweet consiste à répéter le tweet d'un autre utilisateur. Aussi, il est attendu qu'un influenceur voit son contenu plus fréquemment répété par la communauté que le contenu d'un autre utilisateur moins influent.

Etant donné que la fonction de retweet n'est pas une fonctionnalité de Twitter, le service ne fournit pas d'information dans les bases de données qu'il offre publiquement sur le nombre de retweets par individu. De ce fait, nous avons dû collecter « manuellement », individu par individu, le nombre de retweets. Etant donné la fastidiosité de cette tâche, nous nous sommes limités à 100 personnes tirées au hasard dans notre échantillon.

Les corrélations montrent sans surprise que le nombre de retweets par individu est fortement corrélé au nombre de suiveurs ($r=,677$), au nombre de tweets ($r=,297$) et de manière intéressante au ratio suiveurs/suivis ($r=,451$). Pour les deux premières corrélations, il apparaît mécanique que plus un individu a de personnes qui le lisent et plus il y a des chances que celles-ci finissent par reprendre son contenu. De la même manière, il paraît évident que plus un individu va produire et plus il va augmenter les chances que son contenu soit lu par ses suiveurs. Il faut donc vérifier que cette mécanique n'épuise pas la relation qui peut exister entre le nombre de retweets et le ratio suiveurs/suivis. En d'autres termes, il faut vérifier que le nombre de suiveurs et de tweets ne médiate pas (au sens de Baron et Kenny, 1986) cette relation. Le tableau 5 montre que la relation « résiste » à la présence de la mécanique induite par le nombre de suiveurs et de tweet et demeure significative ($t=4,749$; $p<,000$). Par ailleurs le modèle explique 59,2% du phénomène et tous les paramètres sauf la constante sont significatifs ($F(3,88)=41,086$; $p<,000$).

Tableau 5 – Régression multiple (variable dépendante : nombre de retweets)

| Paramètre | Coef. standardisé | t | p |
|-----------|-------------------|-------|------|
| Constante | | ,095 | ,924 |
| Suiveurs | ,562 | 7,702 | ,000 |
| Ratio S/s | ,332 | 4,749 | ,000 |
| Tweets | ,197 | 2,770 | ,007 |

Ce résultat semble donc confirmer que le ratio S/s est un bon indicateur du potentiel de l'individu à être un influenceur. Dans les lignes qui suivent, nous allons approfondir ces premiers résultats en nous focalisant sur l'analyse des contenus produits par des individus présentant un fort ratio suiveurs/suivis par rapport au reste de la population.

3.4 – Contenu produit par les influenceurs sur Twitter

Sur la base du ratio suiveurs/suivis que nous pensons être un indicateur pertinent du leadership des individus sur Twitter, nous avons créé trois groupes de personnes. Un groupe d'utilisateurs présentant un fort ratio suiveurs/suivis qui sur la base de ce qui a été dit plus haut seront considérés comme les influenceurs. Un deuxième groupe constitué de connecteurs ayant un ratio S/s « normal » proche de 1. Enfin, nous avons constitué un troisième groupe de personnes constitué d'individus ayant un faible ratio S/s, inférieur à 0,3 et étant *a priori* considérés comme des suiveurs c'est à dire des personnes en recherche d'information. Enfin, pour éviter les ratios uniquement dûs à un très faible nombre de liens sociaux (.e.g 5 suiveurs et 1 suivi), un nombre minimum de 200 suiveurs et/ou suivis a été fixé pour chacun des utilisateurs.

Etant donné le grand nombre de phrases (tweets) produites par notre échantillon durant deux mois, nous avons sélectionné 30 personnes par groupe et nous sommes limités à l'analyse de 100 derniers tweets, ceci dans un souci de comparabilité de la production. En effet, si il est aisé de trouver des individus avec un fort ratio S/s ayant produit plus de cent tweets en deux mois, l'exercice est plus délicat pour les suiveurs. Heureusement Twitter permet de remonter sur les contenus produits avant la date de suivi. Ceci nous a permis de constituer un groupe 3 de 30 personnes ayant produits au moins une centaine de tweets en remontant de deux mois supplémentaires pour certains. Au terme de la collecte des tweets, nous avons 9254 phrases à analyser (3114 dans le groupe 1, 3135 dans le groupe 2 et 3005 dans le groupe 3).

Caractéristiques « techniques » de la production de contenu.

Avant de nous intéresser au contenu des tweets, une première analyse comparative des individus des différents groupes vient confirmer que le ratio S/s est un bon élément discriminant sur la population des utilisateurs de Twitter.

Tableau 6 – Proportion de hashtags, liens, discussions et logiciels

| Paramètre | Influenceurs | Connecteurs | Suiveurs | Total | Khi ² (2ddl) | P |
|------------|--------------|-------------|----------|-------|-------------------------|------|
| hashtag | 8,4 % | 10,1% | 13,9% | 10,7% | 43,08 | ,000 |
| liens | 29,5% | 23,6% | 38,8% | 30,5% | 147,4 | ,000 |
| discussion | 42,4% | 55,9% | 33,2% | 44% | 286,3 | ,000 |
| logiciel | 56,7% | 80,5% | 60,4% | 60,4% | 457,39 | ,000 |

L'analyse n'a pas été conduite sur l'ensemble des Tweets. En effet, parmi ceux-ci figurent des retweets, c'est-à-dire des contenus qui ne sont pas produits par l'individu. Aussi, même si le contenu des retweets paraît très intéressant d'un point de vue managérial puisqu'il renseigne sur les contenus qui provoquent le plus de médiatisation, il ne sera pas traité ici mais lors d'une étude ultérieure. Le propos étant ici de se focaliser sur ce qui différencie les influenceurs du reste de la population, les retweets ont donc été retirés de l'analyse pour permettre de se focaliser par la production réelle de chaque utilisateur. Il y avait en tout 1097 retweets dans notre échantillon. Une fois ces éléments ôtés il restait donc 8157 tweets à analyser. Il est à noter que l'utilisation des retweets d'un groupe à l'autre est significativement différente ($Khi^2(2)=20,735$; $p <,000$), de telle sorte que de manière attendue les suiveurs (faible ratio S/s : 13,6% de retweets) y ont davantage recours que les « connecteurs » (ratio S/s proche de 1 : 12,2% de retweets) et les influenceurs (fort ratio S/s) sont ceux qui y ont le moins recours (9,9%). Ceci signifie donc que les influenceurs produisent davantage de contenu original que les autres utilisateurs.

Il apparaît tout d'abord de ces résultats (cf. tableau 6) que sur la base du ratio S/s, les caractéristiques des messages produits par les individus sont significativement différentes. Nous nous sommes particulièrement intéressés à quatre caractéristiques.

1 – Hashtag : Ainsi que nous l'avons expliqué dans la première partie de ce travail, le hashtag est une catégorie destinée à organiser les contenus de manière à pouvoir en identifier

rapidement les thématiques. Il ressort des données que le ratio S/s est inversement corrélé à l'utilisation de hashtags. Ceci peut paraître cohérent. En effet, si l'on en croit la littérature, le leader d'opinion possède en général un degré d'expertise sur lequel il s'exprime facilement (Gilly *et al.*, 1998). Etant donné qu'il produit le contenu, ce n'est pas lui qui le classe. Inversement à l'autre extrémité, le suiveur est un consommateur d'information. Il suit activement un flux d'informations qui doivent être classées pour avoir un sens. Il n'est dès lors pas étonnant que cette structuration ressorte lors de ses propres productions. Quoique cette explication soit empreinte d'une certaine logique, elle n'est pas cohérente avec l'approche communautaire selon laquelle le leader socialise et structure sa communauté (e.g. Schouten et McAlexander, 1995)

2 – Liens : Un tiers des messages produits font référence à des informations extérieures à Twitter. Par le biais de lien HTML, l'émetteur du contenu invite ses suiveurs à se rendre sur un site web pour trouver de l'information. Les données montrent que l'influenceur reste dans la moyenne de la population alors que les suiveurs en fournissent beaucoup plus. Ceci est cohérent également avec le fait que le degré d'expertise de l'influenceur peut éventuellement lui permettre de produire un contenu qui est complet quant aux informations à transmettre alors que les suiveurs ont besoin de s'appuyer davantage sur des sources externes. Ceci ne sont bien évidemment que des conjectures et des analyses plus fines doivent être conduites pour mieux comprendre le phénomène.

3 – Discussion : De manière assez évidente, les messages faisant partie d'échanges avec d'autres utilisateurs sont moins fréquents chez les suiveurs qui sont ceux qui ont le moins d'interactions possibles et inversement élevés chez les connecteurs qui sont ceux qui ont le plus grand nombre de relations (beaucoup de suiveurs, beaucoup de suivis). Les influenceurs se trouvent entre les deux du fait qu'ils ont un grand nombre de suiveurs qui peuvent s'adresser à eux. Cette relation apparaît donc plutôt mécanique et due au fonctionnement même de Twitter mais la section suivante nous en apprendra davantage sur ce point.

4 – Interface. La prolifération de logiciels de gestion des tweets et des suiveurs et suivis pose l'utilisateur devant le choix d'utiliser l'interface officielle du Twitter (site Web) ou d'utiliser un logiciel spécialisé. Nous pouvons remarquer que les influenceurs sont ceux qui utilisent l'interface officielle car ils produisent avant tout du contenu sans se soucier de ce qui est produit par les autres. En effet, les logiciels tels que TweetDeck (le plus utilisé par notre

échantillon) permettent de visualiser de manière plus ergonomique tout ce qui est produit par les autres utilisateurs et permet en outre un affichage différencié des messages, des retweets et des discussions. Il n'est pas étonnant dans ce contexte que les connecteurs, soucieux de l'équilibre suiveurs/suivis et ayant un grand nombre d'interactions à gérer aient plus facilement recours à ce type d'outil (80,5%). Les suiveurs l'utilisent également davantage que la moyenne ce qui n'est pas étonnant au vu de l'utilisation principalement informationnelle qu'ils font de Twitter.

Contenu des messages produits.

Avant toute chose il est important de noter que le fonctionnement de Twitter provoque deux types de messages : des messages de discussion entre utilisateurs qui renvoient à la fonction de « chat » de Twitter, et des messages improvisés liés à la fonction première de Twitter : le micro-blogging. Dans le cadre d'une étude sur le leadership, il nous semble opportun de traiter isolément chaque type de contenu. Etant donné que le codage des thématiques a été fait à la main, nous nous sommes restreints à 900 échanges dans chaque groupe, soit 2 700 tweets analysés, répartis en 900 (3*300) discussions et 1800 (3*600) contenus spontanés.

En ce qui concerne les discussions, il est à noter que la plupart des échanges sont des échanges de courtoisie. Dans une forte majorité, il s'agit de phrases permettant de saluer, ou discuter : « passe chez moi il reste un peu de vaisselle ! » (suiveur_17), « bienvenue mister SEO! » (influenceur_9). Suivent ensuite des échanges sur diverses thématiques. Il est alors intéressant de voir parmi nos profils (influenceur, connecteur, suiveur) ceux qui sont à l'origine des questions et des réponses.

Tableau 7 – Proportion de messages de courtoisie, et proportion de questions et réponses

| | Influenceurs | Connecteurs | Suiveurs | Total | Khi ² | p |
|------------|--------------|-------------|----------|-------|------------------|------|
| Courtoisie | 70% | 80% | 68% | 73% | 24,97 | ,000 |
| Questions | 13,3% | 40% | 56,25% | 36,5% | 69,32 | ,000 |
| Réponses | 86,7% | 60% | 43,75% | 63,5% | 48,02 | ,000 |

Le tableau 7 suggère que les discussions de courtoisie sont plus fréquentes chez les connecteurs, ceci est cohérent avec les résultats précédents qui suggéraient que cette catégorie est dans un réseau social plus dense que les deux autres catégories. Le maintien de ce réseau

social passe a priori par l'échange de messages qui n'ont pas forcément un autre but que celui de maintenir les liens sociaux.

Conformément à nos attentes, le groupe d'influenceurs, sur les 30% de discussions qui ne sont pas fondées sur des échanges de courtoisie, est producteur de beaucoup plus de réponses (86,7%) que de questionnements. Les thématiques abordées concernent souvent l'informatique « *c'est le cas chez plusieurs registrars pour éviter que le NDD soit en cours de déplacement lors du renouvellement* » (influenceur_5) mais peuvent être également des conseils en finance « *pour éviter de faire un virement transfontalier avec frais de change de devise pour régler la moindre facture* » (influenceur_2), juridique « *un juge comparera les deux sites en se demandant s'il y a copie manifeste et possibilité de confondre les deux* » (influenceur_21), sexuel, vestimentaire, etc. Le groupe des connecteurs est moins marqué mais demeure quand même un groupe davantage orienté vers les réponses que les questions. Ce groupe se distingue toutefois par des thématiques beaucoup moins techniques que le groupe des influenceurs qui sont reconnus par la littérature comme possédant un degré d'expertise supérieur au reste de la population (e.g. Fejlaoui et Vernet, 2009) et donc aux connecteurs (Goldenberg *et al.*, 2009). En effet, même si certaines réponses concernent directement l'informatique ou les nouvelles technologies « *C'est une erreur fonctionnelle. Mets ton soft en plein écran, cale ton curseur en haut à droite, clique : youpi !* » (connecteur_9), la plupart sont sur le tourisme, « *rend toi à Aoste en Isère, tu vas halluciner. Le plus éco c'est le talon que tu tranches toi-même* » (connecteur_18), la décoration, ou la mécanique. Cependant, étant donné le faible nombre de personnes (30) suivies et de contenus produits par groupe (300 tweets de discussion) il paraît nécessaire d'aller plus en avant dans l'exploration des caractéristiques de ce groupe d'utilisateurs. Enfin, de manière attendue, le groupe des suiveurs pose davantage de questions qu'il fournit de réponses. On remarque cependant que la proportion questions (56,25%) réponses (43,75%) est nettement plus équilibré que le groupe des influenceurs. Si on compare la moyenne des trois groupes, il apparaît qu'il y a beaucoup plus de réponses que de questions. Ceci vient du fait que pour chaque question, plusieurs utilisateurs donnent en général leur avis et qu'une seule question génère de fait plusieurs réponses.

Il ressort donc de cette première phase fondée sur l'analyse du rapport questions / réponses que le ratio suiveurs / suivis utilisé pour constituer les groupes demeure un bon indicateur du leadership des utilisateurs de Twitter dans la mesure où il reste cohérent avec le rapport

d'influence qui est attendu de ce type de population (Vernette, 2003). Dans les lignes qui suivent nous allons enfin nous intéresser au type de contenu qui est produit par les utilisateurs en dehors de toute discussion avec le réseau relationnel.

Il s'agit ici d'analyser les contenus qui sont produits par les utilisateurs de manière spontanée, conformément à la fonction première du service Twitter : le micro-blogging. Pour réaliser cette analyse de contenu, nous avons codé les différents tweets en quatre grandes catégories ; (1) consommation, entreprises, management, (2) politique, environnement, juridique, (3) informatique, NTIC, téléphonie, jeux vidéos, et (4) humeur, vie intime, humour, réflexion, musique (à écouter). Cette catégorisation a été faite au contact des données et n'est dictée par aucun cadre théorique particulier. Des études plus précises sont à conduire pour étudier les tweets à la lumière d'une typologie plus fine des messages. La proportion de chaque catégorie de tweets par groupe figure dans le tableau 8.

Tableau 8 – Répartitions des catégories de message selon le groupe d'utilisateurs

| | Influenceurs | Connecteurs | Suiveurs | Total | Khi ² | p |
|-----|--------------|-------------|----------|--------|------------------|------|
| (1) | 21,12% | 21,62% | 5,59% | 16,36% | 80,08 | ,000 |
| (2) | 10,98% | 5,92% | 0,77% | 6,06% | 61,84 | ,000 |
| (3) | 18,30% | 19,85% | 4,50% | 14,44% | 76,08 | ,000 |
| (4) | 49,57% | 52,59% | 89,13% | 63,13% | 275,3 | ,000 |

D'après les proportions affichées dans le tableau 8, il apparaît que le groupe des influenceurs se caractérise principalement par son recours plus fréquent aux sujets liés à l'environnement (politique, naturel ou juridique). En ce qui concerne les autres catégories, ses spécificités sont moins claires à établir tant les proportions de recours aux trois autres catégories sont proches du groupe des connecteurs. En revanche, *a contrario*, le groupe des suiveurs se distingue clairement des autres en montrant un contenu essentiellement intime (89,13% de messages).

L'influenceur semble donc plus que les autres soucieux de donner son avis sur le monde qui l'entoure. En comparaison des autres groupes il hésite moins à afficher ses opinions politiques ou encore à être critique vis-à-vis de l'actualité politique ou juridique (hadopi, taxe carbone, etc.) En ce qui concerne les opinions politiques il s'agit de diriger le lecteur vers des informations partisans. Il en est de même pour les opinions et critiques sur l'actualité juridico-politique : « *Liste des taxes créées depuis l'élection de Sarkozy* :

<http://twurl.nl/mlfyld>, chez le Faucon » (influenceur_28) ou « Elle rêve de Fred Lefebvre mais pétitionne pour Hadopi. La dernière bourde en ligne de l'UMP <http://tinyurl.com/kwcjvl> #comfoireuse.com » (influenceur_4).

Concernant la consommation, le groupe des influenceurs (21,12%) n'est pas très différent du groupe des connecteurs (21,62%) et donc un peu au dessus de la moyenne (16,36%) compte tenu du plus faible niveau du groupe de suiveurs sur ce point (5,59%). Les écarts entre les groupes sont toutefois significatifs ($\chi^2=80,08$; $p<,000$). Pour les influenceurs donc, plus d'un Tweet sur 5 est directement dirigé vers la consommation, nommant parfois expressément des produits ou des marques : « LCD Toshiba 82 cm à 299 euros sur cdiscount, y'a une arnaque qqpart ? : <http://tinyurl.com/mef7v4> » (influenceur_14). Dans la plupart des cas, les tweets sont en rapport avec des opinions, des critiques, des conseils et des informations sur des produits ou des marques « Si vous souhaitez faire de bons achats lors de la prochaine foire aux vins en Septembre, pensez à Ecce vino : <http://tinyurl.com/lcs3ys> » (influenceur_20). Il peut également s'agir de l'écoute d'un album de musique ou du dernier film vu au cinéma « Ce que j'ai pensé d'Inglourious Basterds, le dernier Tarantino ! <http://bit.ly/14VSUN> Et vous ? Vous l'avez vu ? » (influenceur_25). Il est à noter que d'une manière générale, bien que le nombre de caractères soit limité à 140, les phrases sont formulées en langage peu abrégé et les fautes d'orthographe sont rares (elles sont d'ailleurs très souvent relevées par un autre utilisateur). Ce constat rejoint une des caractéristiques des leaders d'opinion sur internet établies par Fejlaoui et Vernet (2009) et également relevée par les professionnels au sujet de Twitter (Vuitton, 2009).

20% des messages concernent l'informatique, les NTIC ou la téléphonie « Le PageRank... Gage de confiance? <http://bit.ly/DxirV> » (influenceur_6).. En cela, le groupe d'influenceurs ne se distingue pas non plus du groupe des connecteurs. Ce phénomène semble principalement lié au fait que Twitter est un service accessible uniquement par le biais d'un ordinateur ou d'un smartphone dont la population d'utilisateurs semble avoir clairement un profil masculin et orienté vers l'informatique, à l'instar de l'internet tout entier à ses débuts dans les années 1990. La démocratisation croissante de l'outil amène à penser que la proportion de discussions techniques va tendre à diminuer. Actuellement, les « trends topics » affichés par le site, sont en général souvent dirigés vers la technologie : Iphone3G, Spotify, Snow Leopard...

Enfin, la moitié (49,57%) des messages produits sont des messages personnels voire intimes. Cette forte proportion de messages personnels est inhérente au fonctionnement de Twitter qui est avant tout présenté à ses utilisateurs comme un service de micro-blogging. Les influenceurs sont le groupe qui en produit le moins, probablement du fait mathématique qu'ils produisent davantage de contenu à plus forte valeur informationnelle par ailleurs (politique, consommation, technique). Ces messages sont liés à l'humeur du moment « *Ah, que ça fait du bien de virer un client* » (influenceur_2) ou de la journée, « *Journée démoralisante au boulot* » (influenceur_8), une note d'humeur « *j'ai hâte d'être à demain pour comprendre la blague de ce midi, elle avait l'air vachement marrante* » (influenceur_13), une réflexion « *Ciel gris, pluie, périph' bouché... Ça doit être la rentrée :)* » (influenceur_5) ou encore une réaction non engagée à l'actualité « *Sim est éternel* » (influenceur_3). On remarque que les suiveurs font un usage tout à fait notable (89,13%) de ce type de message ce qui est tout à fait cohérent avec leur statut de consommateur d'information.

Conclusion.

Cette recherche avait pour objectif premier de proposer un indicateur permettant d'identifier rapidement et de manière peu coûteuse les leaders d'opinion sur internet également appelés « influenceurs » (Verette et Florès, 2004). Les résultats présentés suggèrent que le ratio suiveurs / suivis est un indicateur pertinent qui permet d'identifier une population qui présente les caractéristiques émergeant de la littérature sur les leaders d'opinion. Ce ratio à l'avantage d'être facilement accessible (le nombre de suiveurs et de suivis étant affichés sur le profil de chaque utilisateur) et utilisable sans délai par les entreprises désireuses de s'adresser à cette population. Des sites tels que twopcharts.com faisant chaque semaine le classement des utilisateurs ayant obtenu le plus grand nombre de suiveurs, il est aisé à l'intérieur de ces listes de calculer quels sont les meilleurs influenceurs selon des thématiques précises. Ce travail permet en outre de présenter le service Twitter comme une source pertinente de données secondaires utiles dans d'autres contextes marketing (Galan et Vignolles, 2009).

Ce travail n'est qu'un premier pas et souffre de certaines limites. La première limite est inhérente au fonctionnement de l'internet et est liée au fait que les comportements et communications observables par les données secondaires sont circonscrites au domaine de l'internet. Ce ne sont pas des individus qui sont étudiés mais des avatars matérialisés par des pseudonymes et dont le comportement est contraint par les règles de l'interface (clavier,

souris, etc.). Les développements en matière d'internet mobile et de géolocalisation amèneront avec le temps des solutions à ce problème. En attendant, il paraît important de développer des recherches qui prennent en compte le comportement effectif des utilisateurs en dehors de l'internet. La deuxième limite est conceptuelle. L'approche exposée ici est essentiellement comportementale et tend finalement à assimiler le concept à ses conséquences (Vermette, 2006). Il paraît donc important de faire une étude de convergence entre la méthode proposée ici avec une méthode dont la validité s'appuie davantage sur la compréhension du concept. Enfin, pour des questions de place, mais également parce que l'objet de cette recherche était uniquement l'identification des influenceurs, nous avons volontairement éludé l'analyse de contenu des retweets qui sont les messages qui présentent le plus d'intérêt pour la communauté. Une voie de recherche importante est l'analyse de ces messages afin d'essayer de comprendre ce qui fait sens dans la communauté Twitter et d'établir la relation particulière qu'entretiennent les influenceurs avec ces messages. Est-ce que le message est digne d'intérêt parce qu'il provient de l'influenceur ou est-ce qu'au contraire c'est le statut d'influenceur qui découle de la capacité de l'individu à découvrir un contenu intéressant ?

Références bibliographiques.

Baron R. M. et Kenny D. A. (1986), The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations, *Journal of Personality and Social Psychology*, 51, 6, 1173-1182.

Belvaux B., Marteaux S. (2007), Les recommandations d'internautes comme source d'information. Quel impact sur les entrées des films au cinéma ? *Recherche et Applications en Marketing*, 22, 3, 65-82.

Bertrandias L. (2004), Etude de la convergence entre deux méthodes d'identification des leaders d'opinion, *Actes du XX^{ème} Congrès International de l'Association Française du Marketing*, Saint Malo, CD-ROM.

Cakim I. (2006), Online opinion leaders: a predictive guide for viral marketing campaigns, in J. Kirby and P. Marsden (eds.), *Connected Marketing: the Viral, Buzz and Word-of-mouth Revolution*, Oxford, Butterworth-Heinemann, 107-118.

Cann Y.M. (2009), Twitter, beaucoup de bruit pour rien, *Observatoire des Réseaux Sociaux*, été 2009, <http://bit.ly/tRqHQ>

Fejlaoui Y., Verneette E. (2009), Repérer les leaders d'opinion dans les communautés virtuelles de consommation : une nouvelle échelle de mesure, *Actes du XXV^{ème} Congrès International de l'Association Française du Marketing*, Londres, CD-ROM.

Florès L. (2008), Web 2.0 : Des études ayant du répondant!, *Décisions Marketing*, 50, 79-82.

Galan J.-Ph. (2002), L'analyse des fichiers log pour étudier l'impact de la musique sur le comportement des visiteurs d'un site Web culturel, *Actes du XVIII^{ème} Congrès International de l'Association Française du Marketing*, Deauville, 255-272.

Galan J.-Ph., Verneette E. (2000), Vers une 4^{ème} génération : les études de marché "on-line", *Décisions Marketing*, 19, 39-52.

Galan J.-Ph., Vignolles A. (2009), Twetnographie : utilisation de Twitter pour la recherche en marketing, *Actes des 14^{èmes} Journées de Recherche en Marketing de Bourgogne*, Dijon, CD-ROM.

Gauzente C. (2006), E-marketing et TextMining : une application à l'analyse des opinions de consommateurs sur Internet, *8^{èmes} Journées Internationale d'Analyse Statistique des Données Textuelles (JADT'06)*, Besançon, 465-742.

Gilly M.C., Graham J.L., Wolfinbarger F. Yale L. (1998), A dyadic study of interpersonal information search, *Journal of the Academy of Marketing Science*, 26, 2, 83-100.

Goldenberg J., Han S., Lehmann D. R., Hong J.W. (2009), The Role of Hubs in the Adoption Processes, *Journal of Marketing*, 73, 2, 1-13.

Holmes J.H., Lett J.D. (1977), Product Sampling and Word-of-Mouth, *Journal of Advertising Research*, 17, 3, 35-40.

Katz E., Lazarsfeld P. F. (1955), *Personal Influence : the Part Played by People in the Flow of Mass Communications*, Glencoe: Free Press.

Kozinets R. V. (2002), The Field behind the Screen: Using Netnography for Marketing Research in Online Communities, *Journal of Marketing Research*, 39, 1, 61-72.

Ladwein R. (1995), Catégories cognitives et jugement de typicalité en comportement du consommateur, *Recherche et Applications en Marketing*, 10, 2, 89-100, <http://bit.ly/ruMFH>

Larceneux F. (2007), Buzz et recommandations sur Internet : quels effets sur le box-office ?, *Recherche et Applications en Marketing*, 22, 3, 45-64.

Roxin V., Bernard Y. (2007), Etiquetage collaboratif et nuages de mots : quels apports pour les sites marchands, *6^{ème} Journée Nantaise de Recherche sur le e-Marketing*, <http://bit.ly/4sWtev>

Schouten J. W., McAlexander J. H. (1995), Subcultures of Consumption: An Ethnography of the New Bikers, *Journal of Consumer Research*, 22, 1, 43-61.

Vernette E. (1997), Evaluation de la validité prédictive de la méthode Delphi-leader, *Actes du XIII^{ème} Congrès International de l'Association Française du Marketing*, Toulouse, 987-1010.

Vernette E. (2003), Les nouvelles perspectives du concept de leader d'opinion en marketing : fondements, apports et pistes de recherche, *Actes du XIX^{ème} Congrès International de l'Association Française du Marketing*, Tunis, CD-ROM.

Vernette E. (2006), Une nouvelle vision du leader d'opinion en marketing : une approche phénoménologique, 5^{ème} Congrès sur les Tendances du Marketing, Venise, <http://bit.ly/1yIgzJ>

Vernette E., Florès L. (2004), Communiquer avec les leaders d'opinion en marketing : comment et dans quels médias ?, *Décisions Marketing*, 35, 3, 23-37.

Vernette E., Giannelloni J.-L. (2004), L'auto-évaluation du leadership en marketing : nouvelle investigation psychométrique, *Recherche et Applications en Marketing*, 19, 4, 65-87.

Vuitton T. (2009), La course aux « followers », *Le Monde*, 12 septembre 2009, <http://bit.ly/1280bJ>