

La circulation des nouvelles sur l'internet : l'interconnexion médiatique au service du pluralisme ?

Emmanuel Marty, Annelise Touboul

▶ **To cite this version:**

Emmanuel Marty, Annelise Touboul. La circulation des nouvelles sur l'internet : l'interconnexion médiatique au service du pluralisme ?. Sur le journalisme, 2013, 2 (1), pp.64-76. hal-00924950

HAL Id: hal-00924950
https://hal.science/hal-00924950

Submitted on 7 Jan 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La circulation des nouvelles sur l'internet : l'interconnexion médiatique au service du pluralisme ?

Emmanuel M ARTY
Maître de conférences
Université de Nice Sophia Antipolis
I3M
Emmanuel.marty@unice.fr

Annelise TOUBOUL
Maître de conférences
Université Lumière Lyon 2
ELICO
Annelise.touboul@univ-lyon2.fr

Résumé
L'article consiste en une étude pluri-méthodologique du paysage des actualités en ligne, dans le cadre d'un projet de recherche intitulé *Internet Pluralisme et Redondance de l'Information*. L'étude cherche à repérer, dans le flux de publication comme dans la circulation d'une nouvelle, les éléments de convergence et de reprise, ou au contraire de créativité et de singularisation des différents types de sites d'information (médias en ligne, infomédiaires, sites natifs de l'internet et blogs) à partir de la publication le 8 mars 2011 d'articles sur un sondage donnant Marine Le Pen en tête des intentions de vote pour l'élection présidentielle de 2012. L'article répond essentiellement à deux objectifs. Le premier est d'identifier le volume de production de l'information dédié au sujet dans les différents sites en tenant compte de la circulation de la nouvelle dans sa dimension diachronique. Le second est d'identifier certains standards dans la production des différents sites, par l'étude des sources et signatures, du traitement visuel ainsi que de l'usage des liens hypertextes. Les résultats dessinent des lignes de convergence et de divergence entre les différentes catégories de sites. Là où les médias en ligne, les infomédiaires et dans une moindre mesure certains sites natifs de l'internet (les plus professionnalisés) se positionnent sur des stratégies de productivité, de réactivité et de reprise normalisée de l'information, les blogs et une grande partie des sites natifs de l'internet (ceux tournés vers l'amateurisme et la participativité) optent pour des valeurs fondées sur la singularité, la créativité, le re-traitement de l'information et la mise en avant d'un réseau de sites "amis", au sein de plateformes techniques néanmoins contraignantes du point de vue de l'interface.

Mots-clés: journalisme, internet, circulation des nouvelles, standardisation de l'information, élection présidentielle

Summary:
This article consists in a multi-methodological study of online news landscape, which is part of a research project called *Internet Pluralism and Redundancy of Information*. The study aims to identify, in the publication flow and in the spreading of online news, some elements of convergence or divergence between different categories of websites (online medias, infomediaries, medias born online and blogs) in their publication, on the 8th of March 2011, of a poll about presidential elections, setting far-right leader Marine Le Pen on top of voting intentions. The article essentially aims at two purposes. The first one is to identify the amount of information production dedicated to this topic by the different websites, including the spreading of the news in its diachronic dimension. The second one is to identify some standards in the production of different websites, through the study of mentioned sources and signings, visual elements as well as the use of hyperlinks. Ours results draw convergence and divergence lines between the different categories of websites. Whereas online medias, infomediaries and to a lesser extent certain media born online (the most professionalized ones), are targeted on productivity and reactivity strategies leading to a normalised rerun of information, blogs and most medias born online (those turned to amateurship and participative publication) make the choice of values based on singularity, creativity and re-treating of information, as well as the promotion of their "friends" websites network, inside technical platforms which, however, can be considered as restrictive for websites interfaces.

Keywords: journalism, internet, news spread, information standardizing, presidential election

Resumo:

Palavra Chave: jornalismo, internet, propagação da notícia, normalização da informação, eleição presidencial

O artigo consiste em um estudo pluri-metodológico do cenário da produção de informação sobre a atualidade na internet, no âmbito de um projeto de pesquisa intitulado Pluralismo na Internet e Redundância da Informação. O estudo busca descobrir, tanto no fluxo de publicação como na circulação de uma notícia, os elementos de convergência e de continuidade ou, ao contrário, de criatividade e de singularização dos diferentes tipos de site de informação (versões on-line de mídias tradicionais, agregadores de conteúdos, sites de notícias nascidos com a internet e blogs) por meio da publicação, em 8 de março de 2011, de matérias sobre uma pesquisa de opinião que colocava Marine Le Pen à frente das intenções de votos na eleição presidencial de 2012 na França. O artigo pretende responder a dois objetivos. O primeiro é idenfificar o volume de produção da informação dedicada ao assunto nos diferentes sites, levando em conta a circulação da notícia em sua dimensão diacrônica. O segundo é identificar certos padrões na produção de diferentes sites, por meio do estudo das fontes e da autoria dos textos, do tratamento visual, bem como o uso de links hipertextuais. Os resultados apontam para linhas de convergência e divergência entre diferentes categorias de site. De um lado, os veículos on-line, os agregadores de conteúdo e, em menor escala, alguns sites nativos da internet (os mais profissionalizados) se utilizam de estratégias de produtividade, de reatividade e de continuidade normalizada da

informação. Já os blogs e grande parte dos sites nativos da internet (aqueles mais ligados ao amadorismo e à participação) optam por valores fundamentados na singularidade, na criatividade, no re-tratamento da informação e na utilização de uma rede de sites "amigos", por meio de plataformas técnicas mas que também possuem constrangimentos do ponto de vista da interface.

La circulation des nouvelles sur l'internet : l'interconnexion médiatique au service du pluralisme ?

Une analyse pluri-méthodologique pour un double questionnement

L'étude que nous souhaitons présenter s'inscrit dans le projet de recherche Internet Pluralisme et Redondance de l'Information (Ipri)[1]. Au sein de cette vaste étude de l'agenda médiatique en ligne, nous avons conduit une analyse pluri-méthodologique de la publication par les différents sites web d'information français d'un sondage donnant Marine Le Pen en tête des intentions de vote pour la présidentielle de 2012. L'objectif est de repérer, dans le flux de publication comme dans la circulation d'une nouvelle, les éléments de convergence et de reprise, ou au contraire de créativité et de singularisation des différents types de sites d'information. Plus précisément, il s'agit de dépeindre les dynamiques d'interaction et d'interconnexion effectives chez les acteurs de l'information en ligne, qui pourraient se manifester différemment d'un type de site à l'autre. Par *type* de site nous entendons des modalités hétérogènes de production et de diffusion de l'information, pouvant être regroupées en quatre catégories, lesquelles sont issues d'une mise à jour (Marty et al. 2012) de précédentes classifications (Deuze 2003, Rebillard 2006, Marty et al. 2010). La première catégorie est celle des médias en ligne, qui recouvre les déclinaisons web des médias pré-existants à l'internet, que ceux-ci soient de presse écrite (*Le Monde*, *Le Point*, etc.) ou audiovisuelle (*RMC*, *France 2*, etc.) La seconde catégorie est celle des infomédiaires, qui rassemble les portails d'information (p.e. *Orange Actualités*) et les agrégateurs de nouvelles (p.e. *Google Actualités*), dont la différence est aujourd'hui ténue. Les blogs (p.e. *Sarkofrance*, *Arnaud Mouillard*, etc.) constituent la troisième catégorie de sites, se distinguant par leur caractère de publication individuelle et autoritative (Rebillard, op.cit.). Enfin, la quatrième et dernière catégorie comprend l'ensemble des sites d'information natifs de l'internet, tels que *Mediapart*, *Agoravox*, etc.

L'enjeu du présent travail est de porter l'attention sur les dynamiques de publication des différents sites, plus que sur la nature de leur traitement du sujet[2], avec l'ambition d'appréhender le processus de publication plutôt que son résultat, dans les dimensions spatio-temporelles et interactionnelles de la circulation du flux de l'information en ligne. L'étude a été conduite pour répondre essentiellement à deux questionnements, distincts mais complémentaires : le premier concerne la circulation diachronique de l'information, incluant l'identification des plages

[1] Programme de recherche ipri - Internet, pluralisme et redondance de l'information (anr-09-jcjc-0125-01b), soutenu par l'Agence nationale de la recherche et regroupant des laboratoires en information-communication (CIM, Université Paris 3 ; ELICO, Université de Lyon ; LERASS, Université Toulouse 3 ; CRAPE, Université Rennes 1 ; GRICIS, Uqam Montréal) et en informatique (LIRIS, Insa Lyon).

[2] Dans un précédent travail (Touboul et al., 2012), nous nous sommes attachés à analyser, par une approche qualitative du discours, le traitement effectué par ces différents sites du sondage sur Marine Le Pen. Nous y questionnons les liens entre sondages politiques, enjeux socio-économiques du web et configurations énonciatives dans le traitement du sujet.

horaires de productivité, la reprise de titres et l'inter-citation de sites, le second est relatif à la standardisation de l'information, à travers l'étude des sources et signatures, du traitement visuel ainsi que de l'usage des liens hypertextes.
Pour ce qui est de la dimension diachronique de la circulation du flux d'information, nous étudierons l'existence de formes normées, "ritualisées" et/ou redondantes dans la diffusion des nouvelles, que l'on peut rapprocher de la notion de "circulation circulaire de l'information" popularisée par Bourdieu (1996) et qui se trouve facilitée sur le web. Nous posons cependant l'hypothèse de l'existence de modes différenciés dans la circulation des nouvelles qui recouperaient largement les catégories de sites précitées. On sait en effet que tous les médias n'ont pas nécessairement les mêmes impératifs en termes de réactivité et/ou d'occupation de l'espace médiatique numérique, en fonction à la fois de l'organisation socio-économique des rédactions et des routines de travail qui en découlent (Lemieux 2000; Neveu 2001; Damian-Gaillard et al. 2009), du contrat de communication proposé aux lecteurs (Charaudeau 2006), ou encore de l'usage que ces derniers font des actualités en ligne (Granjon & Le Foulgoc, 2010). Là où l'internet constitue pour une majorité de sites le lieu du flux et de l'immédiateté, certains sites tels que *Mediapart*, par exemple, sont apparus dans de précédentes recherches comme "moins soumis aux pressions et logiques économiques, aux temporalités imposées par l'AFP et la pression du scoop" (Canu & Datchary, 2010: 209). Il nous appartient donc de décrire de manière diachronique le panorama de la publication en ligne d'une nouvelle, pour préciser cette articulation entre temporalités de publication et cadres de production. Notre seconde hypothèse concerne l'établissement de normes stabilisées dans le domaine de l'inter-citation et de l'interconnexion des différents sites d'actualité, une nouvelle fois adossées aux catégories de sites. L'omniprésence des signatures d'agences de presse, les mentions ou citations d'autres sources médiatiques et l'utilisation des liens hypertextes pour afficher des relations de proximité avec d'autres sites sont autant d'éléments d'inter-connectivité des sites, qui construisent des territoires distincts dans l'espace médiatique numérique. Ces territoires peuvent être marqués soit par une standardisation des procédés de publication et une normalisation de la référence aux agences de presse, soit plutôt par le partage d'opinions en amont de la production des médias de masse et en aval de cette dernière par la pratique de citations réciproques. Dans le domaine politique, Serfaty (2006) et Fouetillou (2008) ont notamment montré, l'une par une étude de la blogosphère politique américaine durant la campagne présidentielle de 2004, l'autre par une étude des blogs français opposés au Traité Constitutionnel Européen en 2005, comment les blogs pouvaient constituer des lieux de médiation supplémentaire, dans lesquels s'expriment des sociabilités fondées sur des logiques partisanes, sinon militantes.

Afin de confronter nos hypothèses à la réalité de la production de nouvelles sur l'internet, il était nécessaire de constituer un corpus exhaustif d'articles produits par les sites web d'information. Dans le cadre du projet IPRI, un logiciel, nommé IPRI News Analyzer, a été développé par une

équipe de programmeurs[3]. Ce logiciel, fonctionnant sur le principe de l'agrégation de flux RSS, nous a permis de collecter plus de 37 569 articles diffusés par 199 sites de différentes catégories[4], entre le 7 et le 17 mars 2011. Plus qu'une simple agrégation, le logiciel permet un archivage des contenus pris en compte en indiquant, pour chaque publication RSS, le site émetteur, la date et l'heure précise de publication, le titre de l'article publié ainsi qu'une partie de son contenu, variable selon les choix opérés en amont par les sites et pouvant aller du simple chapeau à l'intégralité de l'article. En outre, une fonctionnalité lexicométrique permet d'identifier les segments répétés dans les différents titres, synonymes de reprise à l'identique de l'un deux, et des tris sont également possibles selon plusieurs entrées, parmi lesquelles la chronologie de publication des flux (voir *infra*).
Nous nous focaliserons ici de manière plus précise sur un jour et un sujet particuliers. Le 8 mars 2011, paraissait un sondage donnant pour la seconde fois Marine Le Pen en tête des intentions de vote pour le premier tour de la présidentielle de 2012. Ce sujet était alors celui qui avait donné lieu à la plus forte production d'articles sur le web avec 466 articles annoncés dans les flux RSS de 81 sites différents pour la journée du 8 mars. Les modalités du traitement médiatique de la publication de ce sondage seront donc l'objet de notre attention dans le présent travail.

La productivité des sites et les temporalités de diffusion de l'information en ligne

Une production quantitativement différenciée des sites

La première question à laquelle nous nous proposons de répondre est celle des modes de publication des sites d'information en ligne, et de leurs liens avec des cadres de production différenciés. La disparité dans le nombre d'articles publiés par les sites-sur le sujet est un premier élément sur lequel il nous semble important de nous arrêter. Lorsque l'on prend en compte la répartition par catégories de sites du volume d'articles produits sur le sujet dans la journée du 8 mars (voir Tableau 1), cette disparité est flagrante: les médias en ligne représentent la moitié du volume d'articles produits, les infomédiaires un quart, le quart restant se partageant de manière encore inégale entre sites natifs de l'internet et blogs, lesquels représentaient pourtant la moitié du corpus de sites avec 102 blogs pris en compte. Ce chiffre confirme des niveaux de productivité extrêmement différents d'un type de site à l'autre sur ce sujet précis, allant de 8 articles ou plus pour chaque infomédiaire à une moyenne d'un billet pour deux blogs environ.

	Médias en ligne (43 sites)	Infomédiaires (14 sites)	Sites natifs de l'internet	Blogs (102

[3] Samuel Gesche et Elöd Egyed-Zsigmond, INSA Lyon. Le logiciel est disponible selon les termes de la licence Creative Commons: http://liris.cnrs.fr/ipri/pmwiki/index.php?n=Public.IpriNA

[4] Ce chiffre correspond à l'ensemble des sites d'information politique et générale ayant publié sur la période donnée.

			(40 sites)	sites)
Proportion d'articles publiés (466 articles au total)	50 %	26 %	15 %	9 %

Tableau 1: *répartition de la production d'articles (flux RSS) par catégorie de sites sur le sujet Sondage Le Pen du 8 mars 2011.*

On observe donc un niveau de publication assez élevé de la part des médias en ligne, faisant écho aux résultats de précédentes recherches sur le journalisme de "desk" dans les rédactions web (Estienne 2007), plutôt focalisé sur la productivité des rédacteurs et s'appuyant largement sur la publication et/ou la réécriture de dépêches d'agence. La production d'articles est également importante chez les infomédiaires, lesquels n'ont pas de production de contenu propre et ancrent leur activité dans une stratégie de rediffusion des contenus produits par d'autres sites, au premier rang desquels figurent les médias en ligne (Smyrnaios & Rebillard, 2010). En reproduisant les contenus des médias en ligne, qui plus est de manière très souvent automatisée, les infomédiaires se mettent dans une position d'assujettissement aux cadres de production de ces sites. Ils misent alors sur l'exhaustivité, la réactivité et l'accessibilité de l'information comme autant d'éléments d'attractivité pour l'internaute. A l'inverse, la production est quantitativement moindre pour les médias natifs de l'internet et *a fortiori* pour les blogs. Ces différences correspondent à des identités et des stratégies différenciées qu'il s'agit à présent de préciser. Pour ce faire, nous nous sommes d'abord penchés sur les sites les plus productifs sur le sondage donnant Marine Le Pen en tête des intentions de vote à la présidentielle de 2012. Il s'agissait d'une part d'identifier de manière plus détaillée, parmi les médias présents sur la toile, lesquels se situaient à ce fort niveau de production, et d'autre part de comparer ce niveau à leur production globale d'articles sur la journée du 8 mars, afin de déterminer le poids du sujet pour chacun des sites.

Site Web	**Nombre d'articles publiés sur le sujet "sondage Le Pen"**	**Nombre total d'articles publiés le 8 mars 2011**	**Proportion de la production dédiée au sujet "sondage Le Pen"**
RTL	44	106	**41.5 %**
Wikio Actualités	33	206	16 %
Le Nouvel Observateur	25	90	**27.8 %**
Google Actualités	22	84	**26.2 %**

France 2	18	88	20.5 %
France 3	18	89	20.2 %
France Soir	17	176	9.7 %
Le Figaro	16	115	13.9 %
Le Journal du Dimanche	16	161	9.9 %
Mediapart	13	52	**25 %**
TV5 Monde	13	102	12.7 %
Voila Actualités	13	94	13.8 %
Newspeg	12	211	**5.7 %**
Orange Actualités	12	86	14 %
Actu2424	11	54	20.4 %
MSN Actualités	11	86	12.8 %
20 minutes	10	106	9.4 %
Agoravox	9	37	**24.3 %**
François Desouche	9	46	19.6 %
Free Actualités	9	186	**4.8 %**

Tableau 2: *poids relatif du sujet "Sondage Le Pen" dans la production globale pour les 20 sites ayant produit un grand nombre d'articles sur le sujet.*

On observe dans le tableau ci-dessus que le site de *RTL France* est de loin celui qui a consacré la plus large part de sa production de flux RSS à ce sujet, avec 41.5 % des articles, suivi du *Nouvel Observateur*, de *Google Actualités* et des sites natifs de l'internet *Mediapart* et *Agoravox*, qui lui ont consacré environ un quart de leur production respective du jour. A l'inverse, les infomédiaires *Free Actualités* et *Newspeg* n'ont consacré qu'une part minime de leur production à ce sujet. Plusieurs interprétations peuvent être données à ces résultats. Une distinction semble tout d'abord devoir être opérée entre médias en ligne et sites natifs de l'internet d'un côté, infomédiaires de l'autre. Là où les premiers opèrent une sélection "manuelle" de l'agenda du jour, les sujets diffusés sont plutôt le fait d'un référencement automatisé d'articles pour les

seconds. On peut alors penser que l'attention particulière des journalistes aux sondages de manière générale (Champagne, 1990), et à celui-ci en particulier, a pu conduire à une domination du sujet dans l'agenda de ces catégories de sites. Du côté des infomédiaires, en revanche, l'exigence d'exhaustivité des sujets mis en agenda et le caractère automatisé de la sélection ont pu aboutir à attribuer une place moindre au sondage dans la masse des articles traités. *Google Actualités* semble toutefois déroger à la règle en réservant un quart de ses articles au sujet. Ce résultat nous amène d'abord à préciser qu'une distinction existe entre portails (tels que *MSN Actualités*, *Orange Actualités* ou *Voilà Actualités*) et agrégateurs de nouvelles, dont *Google Actualités* fait partie. Là où les premiers puisent leur contenu d'actualité essentiellement chez les agences de presse, qu'ils rétribuent pour leur contenu, les seconds effectuent plutôt une recension de l'ensemble des médias en ligne, dont on a vu qu'ils avaient privilégié le sujet du sondage. A l'intérieur même des agrégateurs, ensuite, la spécificité de *Google Actualités* réside dans sa volonté de "coller" à l'agenda des principaux médias en ligne, y compris dans leur dimension d'analyse et d'opinion liée à l'actualité. Là où des agrégateurs comme *Free Actualités* ou *Wikio Actualités* ont une politique d'agrégation très exhaustive et assez peu hiérarchisée, *Google Actualités* a au contraire mis en œuvre les procédés d'une "traduction algorithmique des logiques éditoriales" des grands médias en ligne (voir Smyrnaios, Rebillard; 2009 pour une présentation détaillée), pouvant expliquer sa proximité avec ces derniers dans nos résultats.

A l'inverse, 33 sites n'ont produit qu'un seul article sur le sujet. En moyenne, il apparaît que ces sites ont produit autour de 6 articles chacun ce jour-là. Parmi ces 33 sites, 24 sont des blogs et 5 sont natifs de l'internet. Plus important, 17 sites, tous des blogs (parmi lesquels *Comité de Salut Public*, *Hashtable* ou encore *La Chronique du Yéti*), n'ont publié qu'un seul article ce jour-là, lequel est donc consacré au sondage sur Marine Le Pen. On voit alors se dégager des profils éditoriaux marqués, allant d'un balayage transversal de l'actualité à une très forte sélectivité des sujets. Le grand nombre d'articles publiés par les infomédiaires doit alors être considéré non comme une attention particulière à ce sujet précis, mais plutôt comme faisant partie d'une stratégie générale d'occupation de l'espace en ligne, exigeant de ces sites à la fois productivité et exhaustivité dans le nombre de sujets traités, comme en atteste la faible proportion d'articles diffusés sur le sondage comparativement au volume global produit. En revanche l'appétence de la blogosphère française pour les sujets politiques, voire polémiques (Cardon 2010), y est évidente, de même que celle d'une grande partie des sites natifs de l'internet. Il semblerait donc qu'à la réactivité et l'exhaustivité des infomédiaires et des médias en ligne, précédemment évoquée, répondent chez les blogueurs et dans une moindre mesure chez les médias natifs de l'internet, une volonté de réflexivité et de singularité, ainsi qu'une productivité largement moindre, exigeant une sélection stricte des sujets jugés dignes d'être traités. Du côté des médias en ligne, seuls certains sites se distinguent par une attention prononcée à ce sujet, parmi lesquels *RTL* et *Le Nouvel Observateur*. Reste néanmoins à déterminer à quelle stratégie éditoriale répond cette forte production d'articles sur le sujet de la part des différents sites. Replacer la production de ces sites dans une perspective diachronique constitue alors un premier élément de caractérisation de cette stratégie. La dimension diachronique va en effet nous permettre de

distinguer les stratégies de réactivité ou de scoop, de modes de fonctionnement tentant de s'affranchir de ces contraintes temporelles, éventuellement d'identifier des cadences de publication ou encore des logiques de suivisme.

Des rythmes de publication marqués

Pour identifier de tels éléments, le logiciel IPRI-NA a systématiquement pris en compte le moment de publication des articles dans le flux RSS des sites, à la seconde près, nous permettant d'établir un véritable film de leur diffusion lors de la journée du 8 mars 2011. Le tableau ci-dessous retrace ainsi le nombre d'articles publiés sur le sujet par les différentes catégories de sites, au cours des différentes tranches horaires de la journée.

	Médias en ligne	**Infomédiaires**	**Sites natifs de l'internet**	**Blogs**	**Total**
0H-3H	0	5	1	1	7
3H-6H	0	1	0	1	2
6H-8H	28	19	2	1	50
8H-10H	35	22	13	3	73
10H-12H	33	17	9	7	66
12H-14H	41	19	6	6	72
14H-16H	24	11	7	3	45
16H-18H	27	14	2	6	49
18H-20H	39	14	4	4	61
20H-24H	21	10	2	8	41
TOTAL	248	132	46	40	466

Tableau 3: *nombre d'articles publiés par catégorie de sites et par tranche horaire*[5]

Plusieurs tendances nettes se dégagent de ce résultat, accréditant l'idée de stratégies de publication différenciées de la part des différentes catégories de sites. On observe tout d'abord

[5] Dans le tableau, les cases en bleu désignent la première tranche horaire de la journée contenant une production quantitativement substantielle, les cases en jaune celles durant lesquelles la production est la plus importante.

que la première tranche horaire à avoir une production d'articles quantitativement significative est celle de 6 heures à 8 heures. Elle ne concerne cependant que les médias en ligne, quasi exclusivement professionnels, ainsi que les infomédiaires, lesquels republient de manière réactive la production des premiers. En ce qui concerne les médias natifs de l'internet, la première plage horaire affichant une production relativement conséquente est plus tardive (entre 8 heures et 10 heures), plus tardive encore pour les blogs puisqu'il s'agit de la tranche suivante de 10 heures à midi. La globalité de ces résultats rejoint ceux de précédentes recherches s'attachant à la dimension diachronique de la circulation des nouvelles (Leskovec et al. 2009), dans lesquelles des chercheurs américains s'intéressaient aux modes de publication des actualités en ligne au sujet de Sarah Palin. La répartition diachronique de la production que l'on observe dans la présente étude montre clairement la contrainte de temps qui s'impose aux médias traditionnels, dont l'utilisation de l'internet est focalisée sur la réactivité, la primeur de l'information, là où les médias natifs et *a fortiori* les blogs, dont l'identité éditoriale repose plutôt sur l'engagement, la singularité et le travail sur l'écriture, s'octroient un délai plus long de publication de l'information. Le caractère particulièrement tardif de la publication des blogs peut trouver au moins deux explications. La première, déjà évoquée par Serfaty (op. cit.), réside dans la fonction de re-médiation prise en charge par les blogueurs, lesquels ne produisent pas l'information première. Ils sont au contraire dans une position de public de l'information délivrée par les médias en général, information dont ils effectuent dans un second temps un commentaire réflexif, construit le plus souvent dans une visée persuasive ou tout au moins argumentative. La seconde raison tient au statut amateur des mêmes blogueurs, lesquels doivent composer par ailleurs avec le rythme des activités quotidiennes, notamment professionnelles et/ou familiales. A cet égard, on remarquera que la tranche horaire durant laquelle les blogueurs ont été les plus actifs sur le sujet est celle de 20 heures à minuit, soit en dehors des horaires les plus répandus de l'activité professionnelle (voire familiale). Pour finir, on notera qu'à l'inverse, les médias en ligne ont été plus productifs entre midi et 14 heures ainsi qu'entre 18 heures et 20 heures, soit sur les tranches horaires lors desquelles les publics sont le plus disponible pour la consultation des nouvelles en ligne. Là où les plages horaires de production semblent bien issues d'une véritable stratégie éditoriale de la part des sites de médias professionnels, les sites amateurs semblent quant à eux plutôt tributaires de conditions de production matériellement contraignantes (impossibilité de produire l'information première, position de public préalable à la production de contenu, contraintes liées aux activités quotidiennes).

Entre duplication, reprise et retraitement : une circulation très disparate des nouvelles

Au vu de ces premiers éléments, il semble clair que les différents sites constituant l'espace médiatique en ligne obéissent à des logiques éditoriales relativement marquées, liées à leur enjeux et contraintes de production. Mais si ces éléments sont explicites dans les dimensions quantitatives et diachroniques, on est en droit de se demander ce qu'il en est d'un point de vue plus qualitatif, c'est-à-dire en ce qui concerne les stratégies de production de sens mises en œuvre par les différents sites. Dans cette perspective, le logiciel IPRI-NA nous a permis

d'identifier, dans la succession chronologique des différentes publications, les titres exacts publiés par nos différents sites, lesquels permettent d'appréhender plus finement les stratégies éditoriales mises en œuvre (voir Figure 1).



Figure 1: *Extrait de la liste diachronique de publication des titres pour le sujet "sondage Le Pen" (Impression d'écran du logiciel IPRI-NA)*

De cette attention particulière portée aux titres successivement publiés sont apparues trois tendances, relatives à des stratégies de publication différenciées. La première d'entre elles est la republication par un site, à intervalles réguliers, du même titre sur son flux: c'est le cas notamment de *RTL*, dont on a vu qu'il était le site ayant consacré au sujet la plus forte proportion de sa production du jour. Le site publie d'abord à 6h11 un article intitulé "*Sondage: Marine Le Pen en tête quel que soit le candidat PS*". A 6h26, l'article est rediffusé sur le flux avec une légère modification (une virgule après "*en tête*"), puis une nouvelle fois, à l'identique, à 6h32. Par la suite, on comptabilisera 18 autres occurrences du titre à l'identique, jusqu'à 12h12, heure à laquelle un nouvel article va succéder à celui-ci ("*Sondage -Marine Le Pen: les sondés "payés"?"*) republié de la même manière dans l'après-midi. Il s'agit là d'une stratégie de republication régulière de l'article, pouvant correspondre à une mise à jour du contenu de l'article (correction d'une coquille, ajout d'une citation), mais témoignant également d'une volonté d'occupation de l'espace des actualités en ligne par le cadencement de la publication sur le flux RSS. Si cette stratégie de republication régulière est particulièrement saillante pour *RTL France*, on peut toutefois noter que c'est également le cas, dans une moindre mesure, pour *Le*

Nouvel Observateur, *France 2* et *France 3* (pour lesquels les flux sont d'ailleurs identiques), *France Soir*, *Le Figaro*, et plus largement les médias en ligne affichant une diffusion conséquente d'articles. C'est également le cas de *Mediapart*. Ce dernier a été le plus productif de sa catégorie sur le sujet. Mais sur les 13 articles publiés par le site sur le flux RSS, on note en réalité 12 publications du même titre et un article original publié en langue anglaise. Il semblerait donc que le site suive des modalités de publication et d'occupation de la toile similaires à celles des médias en ligne "traditionnels", ce qui semble aller à l'encontre de précédentes observations mentionnant une soumission moindre du site aux temporalités et aux logiques économiques ayant cours sur le web (Canu & Datchary, 2010).

Une seconde tendance observée est la reprise, par plusieurs sites cette fois, d'un même titre à l'identique. C'est logiquement le cas des infomédiaires dont la republication d'articles produits par d'autres est le principe même de l'activité. Cette republication s'effectue alors de manière parfaitement simultanée (c'est notamment le cas de *Alice Actualités* republiant le titre de *TF1* et *LCI*: "Mais qui nourrit le Front National?"), ce qui démontre le caractère entièrement automatisé du procédé de republication sur le flux RSS. Plus étonnant, c'est également le cas pour certains médias en ligne. On notera, à titre illustratif, la publication parfaitement simultanée (7:36:09) par *TV5 Monde* et *France Soir* du même titre *"Pour la première fois Sarkozy éliminé du second tour de la présidentielle"* (titre repris 47 secondes plus tard par *Orange Actualités*). On peut alors penser qu'il s'agit là d'une (re)publication automatique d'un fil d'agence, reçu et redistribué par ces médias en ligne, lesquels jouerait alors à leur tour le rôle d'infomédiaires, en ce qu'ils n'effectuent pas de travail de re-médiation journalistique de l'information "brute".

Dans ce paysage caractérisé par une redondance très marquée, les médias natifs de l'internet et les blogs sont les catégories de sites contribuant majoritairement à la diversité de l'offre d'information en ligne. Il semblerait que le délai de diffusion de l'information précédemment évoqué soit pris au bénéfice d'une re-médiation effective de l'information. En témoigne le premier billet de blog publié par *La Chronique du Yéti* à 01h22 (heure que l'on peut interpréter comme une veille tardive du jour précédent) intitulé "*Notre pays est-il condamné à boire le calice jusqu'à la lie?*", ou le billet plus tardif (10h44) et décalé de *A perdre la raison* intitulé *"Marine Le Pen en tête? On s'en fout!"*. Du côté des sites natifs de l'internet, à l'exception de la stratégie de republication régulière de *Mediapart*, précédemment évoquée, et du site nationaliste *François Desouche* effectuant une revue de presse "manuelle" des articles traitant du sujet, laquelle a abouti à des duplications de titres, les articles publiés sur le flux RSS par les sites de cette catégorie sont des articles uniques.

On voit donc se dessiner des lignes directrices dans la politique éditoriale de publication des articles sur le flux RSS, qui coïncident de manière relativement systématique avec les cadres socio-économiques et techniques de production de l'information propres aux différentes catégories de sites précédemment évoquées. Là où médias en ligne et infomédiaires se définissent avant tout par leur souci de la réactivité et d'une occupation régulière de l'espace des actualités en ligne, les blogs se distinguent par une production quantitativement faible, dans un délai plus grand, et dans un souci de ré-médiation et de singularisation de leur identité par

rapport aux autres sites peuplant la toile. Les médias natifs occupent également, pour une large part, cette position, avec toutefois pour certains d'entre eux des politiques les rapprochant des médias en ligne, signe d'une certaine professionnalisation de l'activité (et des enjeux économiques qu'elle suscite) malgré un mode de publication distribuée largement répandu, mêlant amateurs et professionnels dans la rédaction et la gestion des contenus.

La standardisation de l'information en ligne : l'importance des sources et références partagées

En parallèle à l'étude du corpus de flux RSS, nous avons enregistré manuellement le texte intégral et les pages-écrans[6] des articles publiés sur le sondage le 8 mars 2011. En écartant les republications des flux précédemment évoquées, nous avons constitué un corpus de 161 articles, présentant l'intérêt de nous permettre de travailler sur le texte complet des articles ainsi que sur l'ensemble des éléments visuels et graphiques qui lui sont attenants. Sur ces 161 articles, 71 ont été publiés par les médias en ligne, 31 par les infomédiaires, 21 par les blogs et 38 par les sites natifs de l'internet. À partir de ce corpus resserré mais aussi plus riche, nous avons conduit plusieurs études. La première, prenant appui sur une analyse lexicométrique[7] a permis le repérage dans le texte intégral des mentions explicites d'autres médias. La seconde, réalisée sur les pages-écrans des articles, a mis en évidence l'importance de certaines sources qui fournissent la matière première de l'information et signent textes et images des articles publiés. L'étude a aussi porté sur l'affichage de références partagées, devenu sur certains sites à la fois une norme d'écriture et un élément de définition identitaire, comme ont déjà pu le montrer de précédentes études de nature cartographique portant sur les liens hypertextes (De Maeyer, 2010).

Quand les médias citent d'autres médias

Nous nous sommes donc d'abord intéressés aux mentions explicites d'autres sources médiatiques citées par les auteurs dans le corps de leurs articles. A l'aide d'un outil lexicométrique, on a en effet pu identifier quels médias étaient cités le plus fréquemment, et lesquels faisaient parallèlement des mentions récurrentes de sources externes. Le premier média le plus cité, en tant que commanditaire du sondage, est *Le Parisien*, mentionné 21 fois par *RTL*, 7 fois par *Wikio Actualités*, 5 fois par *Google Actualités*. Il est certain que la publication d'un sondage polémique constitue une stratégie efficace de mise en visibilité d'un titre de presse, en permettant d'apparaitre dans les médias concurrents comme source première de l'information suscitant le débat. Dans la même logique le site *Mediapart* a également été cité de manière récurrente: 13 fois par *RTL*, à nouveau, 12 fois par *France 2* et *France 3* et 3 fois par *TV5 Monde*. Le détail des articles montre que *Mediapart* a effectivement été l'auteur d'une information

[6] L'enregistrement des pages-écrans a été réalisé avec le programme *Screengrab*, http://www.screengrab.org/

[7] A l'aide du logiciel Lexico 3. http://www.tal.univ-paris3.fr/lexico/

première, que l'on peut qualifier de scoop, avec ce que le terme implique de sensationnel. Il fut en effet le premier à annoncer que l'institut de sondage Harris Interactive proposait un jeu concours aux sondés participant à l'enquête, sous forme de bons d'achat. Le site qui, rappelons-le, a republié son article une douzaine de fois sur son flux RSS dans la journée, a alors titré *"Sondage loterie: les sondés de l'institut Harris sont payés!"*, soulevant une polémique qui lui a valu de multiples citations de la part d'autres sites. Ce dernier exemple, en plus de conforter l'hypothèse selon laquelle *Mediapart* se rapprocherait des stratégies de publication en ligne des médias professionnels traditionnels (c'est-à-dire venant d'autres supports que le web), montre bien comment le flux de nouvelles en ligne peut rebondir dans une même journée, du fait l'inter-connectivité de la plupart des sites d'information présents sur l'internet. La question qui se pose à présent est de déterminer quels sont les ressorts possibles de cette inter-connectivité, étant entendu que la publication simultanée de la même information par différents sites, les stratégies de primeur de l'information ou d'alignement sur la concurrence, ou encore la constitution de réseaux d'opinion sciemment construits sur la base d'un refus de l'information dominante relèvent de logiques éditoriales tout à fait différentes qu'il nous appartient à présent de préciser.

Les agences de presse : sources et signataires de l'information

Au-delà de la citation des médias entre eux, la fourniture et la signature par une même instance de nombreux textes et images présentent l'intérêt de mettre en évidence l'origine commune de l'information et la faible remédiation par les sites concernés. La question des signatures d'articles comme d'images se pose de façons distinctes sur les différentes catégories de sites. En effet, si la signature des billets de blogs paraît accessoire puisque les publications sont généralement attribuées au blogueur lui-même, elle reste essentielle pour les journalistes des médias en ligne qui marquent ainsi leur statut d'auteur et les droits afférents. Pour les infomédiaires, il importe essentiellement d'attribuer la responsabilité des contenus à l'éditeur pourvoyeur de l'information, sans accorder d'attention à l'auteur individuel des textes et des images. Enfin, pour les sites collaboratifs, la signature de l'auteur est essentielle puisqu'elle constitue la rétribution symbolique des contributeurs bénévoles.
L'attention portée aux signatures se justifie donc car elles constituent un indicateur du statut du texte et de l'image, de leur niveau d'originalité, de la volonté de leurs auteurs de les revendiquer comme une production auctoriale ou bien à l'inverse comme des éléments sans marqueurs énonciatifs, sans responsabilité. La récurrence de certaines signatures comme celle des agences de presse indique une information partagée, issue d'une source unique à laquelle les médias n'apportent pas ou peu de valeur ajoutée.
Les agences de presse (*AFP, Reuters, Max PPP*, etc.) signent la grande majorité des photographies publiées par les médias en ligne (soit 38 des 65 images publiées) et par les infomédiaires (24 des 43 images publiées). Cette pratique se retrouve aussi chez certains sites natifs de l'internet, notamment ceux qui adoptent pour une part de leur activité, le modèle des médias en ligne. Juste après les agences, les médias audiovisuels apparaissent comme d'importants pourvoyeurs d'images reprises sur l'ensemble des sites d'information, même si

cette pratique est moins répandue sur les blogs. Les chaînes de télévision comme *France Télévision* et *TF1/LCI* diversifient ainsi les canaux de diffusion de leurs productions en devenant une ressource précieuse pour des médias (essentiellement imprimés) qui ne disposent pas de compétences en matière d'images animées. Si l'image ne constitue pas leur cœur de métier, certaines radios, particulièrement *Europe 1*, mettent à la disposition des autres médias les séquences filmées des interviews réalisées en studio. Enfin, la reprise de vidéos publiées dans un premier temps sur *Youtube* ou *Daily Motion* permet à des sites de bénéficier d'images dont le crédit est incrusté à l'écran (généralement des images provenant de médias audiovisuels ou radiophoniques). Ces signatures participent d'une stratégie de marquage du territoire éditorial et de recherche de visibilité déjà identifiée avec l'étude diachronique de publication des flux RSS.

Concernant les blogs et les billets publiés par des sites natifs de l'internet, les images quand il y en a, car l'illustration est nettement moins systématique que dans les médias en ligne et leurs dérivés, se présentent souvent sans signature. Cela peut s'expliquer par le fait que ces images ne sont pas nécessairement issues de l'actualité mais empruntent plutôt à la figure rhétorique de la métaphore. Dans cette perspective, l'image ne sert pas à montrer de façon plus ou moins dénotée ce dont il est question dans l'article mais elle constitue un élément discursif qui contribue à la tonalité générale du billet et à travailler le lien avec les lecteurs. La connivence supposée, liée au partage par le blogueur et ses lecteurs des mêmes opinions, doit permettre à ces derniers de décoder le sens voulu par l'auteur malgré l'absence de signification explicite. A ces résultats généraux, il faut ajouter quelques observations ponctuelles. En dehors des blogs, certains médias n'indiquent aucun crédit sur les images publiées. C'est le cas du *Monde* par exemple, alors même que ces images sont des images d'agences, publiées avec leur crédit sur d'autres sites. À l'inverse, *Le Post* est le seul site à mentionner systématiquement les sources iconographiques et informationnelles. Un espace à renseigner est ainsi prévu sous le texte que les contributeurs occasionnels ou rédacteurs du site remplissent le plus souvent. C'est sur ce site que la distinction entre auteur et sources est la plus explicite.

La signature des articles semble moins systématique que pour les images sur les médias en ligne. Sur les 71 articles de cette catégorie de sites sur le sujet du sondage, seuls 25 sont signés par une agence à laquelle est associé le nom du site éditeur ou celui d'un rédacteur. À cela, il nous semble pertinent d'ajouter les 23 articles non signés dont on peut supposer qu'il s'agit de dépêches d'agence pas ou peu retravaillées, correspondant à l'activité de "bâtonnage" peu gratifiante mais imposée aux journalistes des rédactions web et dont Yannick Estienne (2007) a montré la pratique généralisée. Si on prend la liberté d'ajouter les articles signés par une agence et ceux qui ne portent aucune signature, le total représente plus de 60% des textes publiés sur les médias en ligne, ce qui nous rapproche des 59% d'images créditées aux agences et confirme l'importance exceptionnelle de ces pourvoyeurs d'informations pour les médias *mainstream.*

Si les médias en ligne, notamment audiovisuels, servent parfois de sources en matière d'image, c'est beaucoup plus rare concernant les textes, à deux exceptions près. *Le Post*, site natif de l'internet, semble ainsi faire œuvre de pédagogie en exigeant que soient mentionnées les sources même déjà citées dans le texte, ce qui a pour résultat l'indication, pour un seul billet, de 3 ou 4

sources parmi lesquelles des médias en ligne. Cette information sur les sources est distincte de la mention de l'auteur qui bénéficie d'un espace à part comme c'est le cas dans la plupart des sites qui accordent une place importante aux contributions des internautes. Autre particularité, le site natif de l'internet *François Desouche* se présente comme un panorama de presse, soit une compilation de médias en ligne sur un sujet donné. Ce site copie et publie l'amorce de vidéos, sons ou articles sur le sujet du sondage, indiquant de façon tout à fait explicite le nom du site d'origine et le lien pour y accéder mais sans ajouter aucun texte original ou commentaire. Pour finir, les signatures sont totalement absentes des blogs qui constituent des territoires propres où la signature du propriétaire-auteur apparaît de fait inutile.

Si le rôle prépondérant des agences de presse n'est pas nouveau et déjà analysé en ce qui concerne les médias audiovisuels ou imprimés (Boyd-Barett, Rantanen, 1998), notre observation confirme l'omniprésence des agences dans la production des médias en ligne (Paterson, 2005) et la très faible remédiation de ces derniers. Dans le paysage de l'information sur internet, l'importance des agences est accrue par la position dominante des médias en ligne dont les informations sont reprises par les infomédiaires.

Si la reprise des dépêches d'agence n'est pas cachée par les médias en ligne, on peut supposer que les signatures *AFP, Reuters* ou *AP* ne retiennent pas particulièrement l'attention des lecteurs. A l'inverse, l'image, quelle que soit sa qualité informationnelle capte le regard. Non seulement elle constitue un élément visuel important souvent imposé dans son association avec le texte mais l'iconographie est un acte éditorial discuté et encadré par les règles définies dans les chartes éditoriales. Peut-on tout montrer? Faut-il choisir des images redondantes avec le texte, simples illustrations ou au contraire utiliser l'image comme un complément discursif à l'écrit ? L'image doit-elle se contenter d'enregistrer et renforcer l'effet de réel ou au contraire proposer un regard singulier sur le monde ? Certains médias ont fait de l'image un des éléments distinctifs de leur signature éditoriale comme *Paris Match* ou *Libération.* Mais les contraintes économiques, le rythme de la publication sur internet et la mise à disposition de ressources par les acteurs de l'information eux-mêmes (comme des comptes *Flickr*) changent la donne. Ainsi, l'attention portée au choix des images au sein des articles de notre corpus permet de compléter l'étude des citations et des signatures pour évaluer le degré d'originalité ou à l'inverse l'ampleur de la reprise et de la duplication en matière d'information sur Internet.

Ayant constaté, l'omniprésence des signatures d'agence de presse en matière de photographie sur les sites des médias en ligne, il n'est pas étonnant de retrouver un certain nombre d'images identiques publiées par différents sites aux lignes éditoriales pourtant distinctes voire opposées si on se réfère aux sensibilités politiques des titres. Ainsi, *Le Monde* qui se plaît à revendiquer le statut de journal de référence publie la même photo de Chantal Brunel que *Le Parisien*, la même photo de N. Sarkozy que *Le Parisien,* la même photo de Xavier Bertrand qu'*Europe 1* et la même photographie de Marine Le Pen que *Le Figaro, RMC* et *La Tribune*. Une autre image de M. Le Pen se retrouve publiée à l'identique sur les sites de *20 minutes, France 3, Les Echos, Metro et RFI*. Enfin, certains sites n'hésitent pas à publier la même photographie sur différents articles. Ainsi on retrouve une photographie identique sur deux articles d'*Europe 1* et une autre

image répétée trois fois sur le site de *France 3*. Dans les cas où les représentations de M. Le Pen ne sont pas totalement identiques, elles proviennent néanmoins des mêmes événements couverts par des photographes d'agence. Il s'agit principalement de meetings ou de conférences de presse pour lesquels les chargés de communication de la leader frontiste appliquent à la lettre les préconisations de la communication politique, à savoir une mise en scène qui impose aux photographes la lumière, les couleurs, les symboles que l'on souhaite associer à M. Le Pen et la distance minimale nécessaire pour empêcher les déformations liées à la focale ou au cadrage choisis. La plupart des images montrent donc M. Le Pen dans un rôle d'oratrice, à la tribune, se détachant d'un fond sombre ou des couleurs du drapeau français. Elle est généralement placée au centre de l'image et photographiée en légère contre plongée. Il semble évident que ces sources partagées que sont les images d'agence contribuent à installer durablement une iconographie de M. Le Pen conforme à ses slogans politiques et à la stratégie de son parti : une femme d'action, qui prend la parole et qui se réfère sans cesse aux symboles patriotiques les plus traditionnels.

Si les médias en ligne montrent généralement M. Le Pen quand les articles traitent de M. Le Pen, les blogs développent des stratégies très différentes. Les blogueurs font souvent le choix de la provocation et de la dérision en transformant des images grâce à des photomontages et des retouches. Le ton est généralement impertinent comme en témoigne l'image de M. Le Pen en Schtroumpfissime dans la cour de l'Elysée, ou encore sur le blog *C'est juste histoire de dire* celle d'une jeune fille blonde peu vêtue dont le visage a été remplacé par celui de JM Le Pen. Parfois, les blogueurs ne montrent pas M. Le Pen mais privilégient le registre métaphorique en proposant à l'appui d'un naufrage national annoncé une photo du paquebot France ou la reproduction de la couverture d'un manuel de plongée pour dénoncer une stratégie politique dangereuse. Là encore, la liberté de ton revendiquée joue sur le registre de la provocation ou de l'humour grivois comme en témoigne l'image choisie par *Le carnet de notes de Yann Savidan* qui présente trois oursons en peluche collés les uns aux autres à l'appui du titre "*Même contre DSK, M. Le Pen jouit encore*". De façon moins radicale, certains blogueurs choisissent parfois des images simplement peu avantageuses pour M. Le Pen la montrant par exemple en train de tirer la langue (*À perdre la raison; Rimbus, le blog*). Pour finir, si les blogs ne présentent pas systématiquement d'illustration à l'appui des billets, les images ne sont généralement pas redondantes avec le texte, rarement "neutres" et jamais partagées avec d'autres sites.

L'étude des signatures et des sources partagées sur un même sujet met en évidence le poids des agences de presse au niveau des producteurs traditionnels de l'information journalistique. En apparence, les médias en ligne semblent produire une information relativement homogène, illustrée par les mêmes images et nourrie par les mêmes dépêches d'agence. Cet état de fait semble logiquement résulter de contraintes de productions ainsi que de normes en matière d'écriture et de pratiques professionnelles qui structurent et définissent le champ du journalisme professionnel. Cependant, à partir de ce premier niveau d'observation, il convient de rester prudent et ne pas conclure à l'homogénéité de l'information des médias en ligne que seule une analyse de discours plus fine pourrait affirmer. Enfin, il ne faut pas non plus conclure que seuls les médias en ligne partagent sources et normes d'écriture. Ayant constaté que les blogs semblent

se distinguer radicalement des médias en ligne en matière d'illustration, nous avons souhaité interroger l'existence de normes et références partagées au sein de cette catégorie de sites avec l'étude des liens hypertextes.

Les stratégies réticulaires des blogs

Loin des limites physiques imposées par les pages de papier ou la linéarité du flux audiovisuel, les sites web présentent des territoires construits autour de nombreuses propositions de navigation. Les liens hypertextes structurent le territoire éditorial avec les rubriques, les menus, les renvois sur des informations complémentaires ou plus anciennes. Ce sont ces mêmes liens hypertextes, définis comme "signes-passeurs" par Souchier et al. (2003) qui dessinent les contours de la publication plus ou moins ouverte sur d'autres sites. L'étude des liens et de leur rôle dans la publication (structuration des contenus mais aussi accès aux publicités, sources, partenaires, amis etc.) permet de mettre en lumière les stratégies qui dessinent la dimension réticulaire de l'écriture de l'information en ligne.
Dans le cas des blogs, le territoire éditorial se donne à voir comme un espace sous la seule autorité du blogueur. Il y est généralement le seul auteur même s'il n'hésite pas à citer d'autres médias et publier les commentaires de ses lecteurs. Les blogs sont générés et administrés grâce à des applications qui ne nécessitent pas de connaître les langages de programmation. Les interfaces normalisées qui en résultent et que les blogueurs peuvent éventuellement personnaliser prévoient généralement le découpage de la page en zones aux fonctions clairement identifiées : entête, menu, billets, commentaires et dans la colonne de droite présentation de l'auteur, fil twitter, archives, nuage de tags et surtout la liste des blogs amis. Structurellement donc, une partie des liens qui s'affichent sur les blogs contribuent à dessiner le prolongement du territoire éditorial vers des sites partageant les mêmes goûts, opinions, passions, centres d'intérêts, etc. (MacDougall, 2005). Si certains font un usage modéré de cette liste d'amis, d'autres blogs à la fois militants et soucieux de leur référencement par les moteurs de recherche proposent une longue liste de liens à l'instar des sites *Les échos de la gauchosphère* ou *Le coucou des claviers* qui en affichent une cinquantaine chacun. Dans la même veine, certains sites natifs de l'internet comme *Bellaciao*, placés dans cette catégorie car animés par un collectif mais proches des blogs par leur engagement, encadrent leur page par deux colonnes de liens sur des événements, des publications et des sites amis de chaque côté des billets. En l'absence de hiérarchie et parfois sans organisation ordonnée (pas même alphabétique), cette mise en réseau se rapproche de la structure rhizomatique telle que conceptualisée par Deleuze et Guattari (1972, 1980) plus que de l'arborescence des pages web des médias en ligne. Le principe du rhizome combine la connexion généralisée à l'hétérogénéité, rejette la notion de centre et d'organisation et dessine une carte dynamique de parcours possibles. Loin de la navigation au sein du territoire éditorial des médias en ligne, circonscrit par le site lui-même et ses partenaires économiques (annonceurs, groupe de communication), les blogs dessinent des navigations possibles au sein d'une sphère éditoriale ouverte, construite par la proximité et le partage des opinions.

Considérant désormais non plus la structure de la page mais l'écriture des billets, deux principales stratégies se distinguent. La première, assez répandue, consiste à multiplier les citations autoréférentielles dans son billet (*Les jours et l'ennui de Seb Musset*, *Marc Vasseur,* etc.). Cette stratégie a l'avantage de mettre en perspective d'anciennes prises de position avec l'actualité la plus récente, donnant ainsi au blog l'image d'une pensée solide qui se développe et s'approfondit en permanence mais aussi quelque peu fermée sur elle-même, quasi autarcique. La seconde stratégie au contraire, inscrit dans l'écriture du billet ses sources d'inspiration et invite le lecteur à y faire un tour, à l'image du blogueur du *Coucou des claviers* qui écrit plaçant quatre liens hypertextes concernant des blogs amis dans la phrase suivante "Hier Yann Savidan avait lancé une chaîne avec son billet "Est-ce que tu iras éteindre le feu ?". Il m'avait demandé ainsi qu'à quelques autres, ce que je ferai dans le cas d'un face à face M. le Pen - N. Sarkozy à la présidentielle". Lorsqu'on parcourt le billet de Yann Savidan, il apparaît que les commentaires sont postés par les blogueurs cités dans le billet du *Coucou des claviers*. Micro-mondes ? Homophilie ? Si l'échelle de l'étude que nous avons conduite ne permet pas de tirer des telles conclusions, elle permet cependant d'affirmer que les réseaux de liens croisés inscrits dans les billets et les commentaires sont les marqueurs de maillages qui dessinent des mondes sociaux (Trédan 2010) ou des blogosphères autour du partage des opinions.

Conclusion

Au terme de cette étude, nous pouvons clairement identifier des lignes de fracture entre les différentes catégories de sites d'information. L'observation diachronique de la publication et de la circulation des nouvelles, les citations inter-médiatiques ainsi que le repérage de la prégnance des agences de presse dessinent un premier ensemble de sites partageant stratégies, contraintes économiques et normes professionnelles. Cet ensemble, qui regroupe les médias en ligne et les infomédiaires, se définit par une productivité importante, une réactivité programmée ainsi qu'une information normée par les pratiques professionnelles du journalisme et massivement diffusée. La primeur de l'information, l'occupation continue des multiples espaces de l'internet et la capacité à devenir une source pour être cité par les médias concurrents semblent alors constituer pour cette catégorie de sites autant de priorités, dut-ce être au détriment du pluralisme de l'information. Le deuxième ensemble, essentiellement constitué par les blogs, se caractérise par une activité de publication seconde par rapport à celle des médias en ligne, par conséquent décalée dans le temps mais également dans le ton. Pour ces amateurs, généralement seuls à alimenter leur publication, l'identité éditoriale passe par un effort de singularisation qui s'exprime avec le travail graphique, photographique et rédactionnel. La connectivité sociale contribue à la mise en réseau de publications certes moins massives mais relayées et circulant au sein de groupes ou de grappes de blogs partageant les mêmes opinions et registres d'expression. La reconnaissance du réseau de pairs semble donc primer, pour cette catégorie de sites, sur les enjeux de rentabilité et de visibilité des contenus produits. A l'intersection de ces deux ensembles très distincts se trouve la catégorie des sites natifs de l'internet. Cet ensemble moins homogène que les deux précédents semble occuper une position intermédiaire et hybride dans le paysage de l'information en ligne. En effet, cette catégorie comprend des sites comme *rue 89,*

Mediapart ou *Le Post* qui, par certains aspects, tentent de s'adosser au modèle des médias en ligne dont ils partagent une part de la culture et des contraintes professionnelles. D'autres sites comme *Bellacio* ou *Alterinfo* se rapprochent plutôt des blogs, malgré la dimension collective de la publication, essentiellement du fait de leur caractère amateur, engagé et non-marchand. Ces distinctions qui se font jour parmi les éditeurs d'informations sur le web confirment ainsi l'aspect déterminant des cadres de production des actualités en ligne, dans leurs dimensions temporelles et économiques et dans les pratiques routinières qui en découlent. Mais elles mettent également au jour le rôle joué par la configuration de normes discursives que partagent les divers acteurs de l'information en ligne, professionnels ou non, avec leurs publics respectifs, et qui contribuent pareillement à définir la physionomie de leur production éditoriale et les relations qu'ils envisagent de construire avec l'ensemble des acteurs du dispositif de l'information sur l'internet.

Bibliographie

Bourdieu Pierre, 1996, *Sur la télévision* (suivi de *L'emprise du journalisme*), Paris : Liber.

Boyd-Barett Olivier, Rantanen Terhi (ed.), 1998, *The globalization of news*, Sage Publications

Canu Roland, Datchary Caroline, 2010, "Journalistes et lecteurs-contributeurs sur Mediapart. Des rôles négociés", *Réseaux*, n°160(2-3), pp. 195-223.

Cardon Dominique, 2010, *La démocratie Internet. Promesses et limites*, Seuil, coll. « La république des idées », 102 p.

Champagne Patrick, 1990, *Faire l'opinion : le nouveau jeu politique,* Paris, Minuit, coll. Le sens commun.

Charaudeau, Patrick, 2006, Discours journalistique et positionnements énonciatifs : frontières et dérives [En ligne]. *Semen*, 22. Disponible sur : http://semen.revues.org/index2793.html

Damian-Gaillard Béatrice, Rebillard Franck., Smyrnaios Nikos, 2009, "La production de l'information web: quelles alternatives? Une comparaison entre médias traditionnels et pure-players de l'internet", *New Media and Information: convergences & divergences,* Athens, international conference proceedings, CD-Rom ISBN 978-960-6746-05-5.

Deleuze Gilles, Guattari Félix, 1980, *Capitalisme et Schizophrénie 2. Mille Plateaux*, Paris, Minuit.

Deuze Mark, 2003, « The web and its journalisms: Considering the consequences of different types of newsmedia online », *New Media & Society*, vol. 2, n° 5, pp. 203-230.

Estienne Yannick, 2007, *Le journalisme après internet.* Paris : L'Harmattan. 313 p.

Fouetillou Guilhem, 2008, « Le web et le traité constitutionnel européen. Écologie d'une localité thématique compétitive », *Réseaux*, n° 147, p. 229-257.

Granjon Fabien, Le Foulgoc Aurélien, 2010, « Les usages sociaux de l'actualité. L'expérience médiatique des publics internautes », *Réseaux* n° 160(2-3), pp. 225-253.

Lemieux Cyril, 2000, *Mauvaise presse. Une sociologie compréhensive du travail journalistique et de ses critiques*, Paris, Métailié.

Leskovec Jure, Backstrom Lars., Kleinberg Jon., 2009, « Meme-tracking and the dynamics of the news cycle », *KDD'09 - International Conference on Knowledge Discovery and Data Mining*, Paris, [http://www.cs.cornell.edu/home/kleinber/kdd09-quotes.pdf].

MacDougall Robert, 2005, "Identity, Electronic Ethos, and Blogs. A Technologic Analysis of Symbolic Exchange on the New News Medium", *American Behavioral Scientist*, vol. 49, n°4, pp. 575-599

Maeyer (de) Juliette, 2010, "Mapping the hyperlinked environment of online news. Issues and challenges for the French news sites", in *IAMCR Conference*, en ligne [http://juliettedm.files.wordpress.com/2010/06/demaeyer_mapping-hyperlink-environment.pdf]

Marty Emmanuel, Rebillard Franck, Pouchot Stéphanie, Lafouge Thierry, 2012, "Diversité et concentration de l'information en ligne. Une analyse à grande échelle des sites d'actualité français". *Réseaux*, n°176, pp. 29-74.

Marty Emmanuel, Rebillard Franck, Smyrnaios Nikos, Touboul Annelise, 2010, "Variété et distribution des sujets d'actualité sur l'internet. Une analyse quantitative de l'information en ligne", *Mots. Les langages du politique*, n° 93, ENS Éditions, pp. 107-126.

Neveu Erik, 2001, *Sociologie du journalisme*, Paris, La Découverte.

Paterson Chris, 2005, "News agency dominance in international news in the internet", *Converging media, diverging politics : a political economy of news media in the united states and canada*, edited by D. Skinner, J.R. Compton, M. Gasher, Lexington Books.

Rebillard Franck., 2006, "Du traitement de l'information à son retraitement. La publication de l'information journalistique sur l'internet", *Réseaux*, n° 137, pp. 29-68.

Serfaty Viviane, 2006, « Les blogs et leurs usages politiques lors de la campagne présidentielle de 2004 aux États-Unis », *Mots. Les langages du politique*, n°80, ENS Éditions, pp. 25-35.

Smyrnaios Nikos, Rebillard Franck, 2009, « L'actualité selon Google. L'emprise du principal moteur de recherche sur l'information en ligne », *Communication et langage*s no 160, p. 95-109.

Smyrnaios Nikos, Rebillard Franck, 2010, "Les infomédiaires au cœur de la filière de l'information en ligne. Les cas de Google, Wikio et Paperblog", *Réseaux* n° 160-161, pp.163-194.

Souchier Emmanuël, Jeanneret Yves, Le Marec Joëlle (sous la dir.), 2003, *Lire, écrire, récrire. Objets signes et pratiques des médias informatisé,* Bibliothèque Centre Pompidou, collection Etude et recherche

Touboul Annelise, Damian-Gaillard Béatrice, Marty Emmanuel, 2012, La disparité des modes de traitement journalistiques et des énonciations éditoriales sur le web. Le cas d'un sondage sur Marine Le Pen et la Présidentielle de 2012. *Réseaux*, n°176, pp. 75-106.

Trédan Olivier, 2010, "La construction des mondes sociaux par la pratique de l'auto-publication : retour sur l'expérience de quelques blogueurs", *Communication & langages*, n° 165, pp.73-86