

Tout ça pour ça ? Titres fonctionnels et égocentrisme hypertextuel

Guillaume Sire

Maître de conférences en Sciences de l'Information et de la Communication

Université Toulouse 1 - Capitole

Lerass / Idetcom

Voilà donc vingt ans que le web existe. Vingt ans que nous pouvons mettre en ligne des documents et les relier les uns aux autres grâce à des liens hypertextes. Vingt ans que nous sommes équipés d'ordinateurs de plus en plus puissants et légers, petits, polyvalents, connectés les uns aux autres par l'intermédiaire de modems, de câbles, de routeurs et des fermes de serveurs. Le réseau a tissé parmi nous sa toile physique et logicielle. Les informations sont échangées, stockées et mobilisées plus rapidement et facilement que jamais.

Les premiers observateurs du web furent nombreux à prédire que le journalisme changerait radicalement (Domingo, 2006). Les liens hypertextes permettraient un rapport direct aux sources et une structure non linéaire du récit (De Maeyer, 2013). Nous assisterions au triomphe du multimédia, mêlant au texte l'image, le son et la vidéo. Les citoyens collaboreraient étroitement avec les professionnels de l'information. Libérés des contraintes d'espace, les journalistes donneraient libre cours à leur créativité. Finie la circulation circulaire de l'information, aucun mimétisme, pas d'uniformité. La diversité et le pluralisme triompheraient. Les différents points de vue, même les plus loufoques et les plus extrêmes, seraient exprimés et entendus. Ouf, la démocratie était sauvée.

Les années qui suivirent montrèrent, et rappelèrent, à quel point il est nécessaire de se montrer prudent à l'égard des raisonnements technodéterministes selon lesquels l'existence d'un nouveau moyen de communication implique nécessairement l'avènement d'une meilleure façon de communiquer. Quand c'est nouveau et différent, ce n'est pas toujours mieux. C'est peut-être simplement nouveau et différent. En tout cas cela s'ajoute à l'existant plutôt que de remplacer quoi que ce soit. Certains problèmes sont résolus mais d'autres apparaissent, et d'autres se transforment, et d'autres demeurent intacts. Le web n'a finalement pas constitué la modification radicale que beaucoup annonçaient mais une nouvelle « configuration sociotechnique » en perpétuelle mutation (Rebillard, 2007), au sein de laquelle se mêlent au point de s'indifférencier l'ancien et le nouveau.

Nous avons consacré notre thèse à l'observation des effets tangibles qu'ont sur les pratiques journalistiques et le récit d'actualité ces configurations et ces reconfigurations (Sire, 2015). A partir des résultats d'une observation effectuée en 2012 dans le cadre de cette thèse, et réitérée en 2015 pour cet article, nous pointerons notre focale sur deux pratiques emblématiques : l'écriture des titres et l'usage des liens hypertextes. Nous montrerons que les titres des articles journalistiques mis en ligne sur les sites adossés à des quotidiens ou à des magazines consacrés à l'information politique et générale se ressemblent. D'autre part, nous montrerons que les liens hypertextes sont peu utilisés par les journalistes et servent essentiellement à pointer vers des pages internes créés à des fins de captation de trafic. Nous discuterons des raisons pour lesquelles les journalistes des sites observés ne semblent pas s'être saisis des possibilités de différenciation et de délinéarisation offertes par le web, avant de conclure en soulignant que c'est précisément et paradoxalement parce qu'ils ne se saisissent pas de ces possibilités que ces sites sont les plus visibles sur le web.

Méthode d'observation

Nous avons sélectionné six sites web adossés à des titres de presse imprimée nationaux d'information politique et générale (quotidiens ou magazines) : Lemonde.fr,

Lefigaro.fr, Lepoint.fr, Lexpress.fr, Leparisien.fr, Nouvelobs.com. Ces sites ont été choisis parce qu'ils étaient les plus visités parmi les sites adossés à des titres nationaux de presse imprimée d'information politique et générale payante selon les chiffres de l'OJD en octobre 2012. Pour chacun de ces sites, nous avons prélevé les trois articles qui se trouvaient en tête de la page d'accueil au moment où nous les visitions (en fin d'après-midi), durant quatorze jours consécutifs en octobre 2012, pour un total de 252 articles. Nous avons ensuite codé à la main une dizaine d'informations par article afin de comparer l'utilisation des signes de ponctuation, les liens entrants et sortants, l'usage des balises HTML. Nous avons réitéré l'observation, de la même façon exactement, en septembre 2015. Nous avons à nouveau prélevé et analysé 252 articles sur les mêmes six sites d'information. Il s'agissait toujours à cette date des sites les plus visités parmi les sites adossés à des titres nationaux de presse imprimée d'information politique et générale payante selon les chiffres de l'OJD.

Notre méthode présente plusieurs défauts dont les principaux peuvent être résumés par les questions suivantes : Pourquoi ne pas avoir choisi d'autres sites de presse en ligne ? Pourquoi avoir exclu les sites non-adossés à des titres de presse papier ? Pourquoi ne pas avoir sélectionné plus d'articles ? Il nous semble toutefois que, malgré les imperfections évidentes de notre méthode, notre observation n'est pas dénuée de qualités heuristiques. Les sites choisis, étant donné leurs moyens de production et leurs modèles économiques, sont comparables. Nous verrons que le nombre d'articles prélevés a permis d'identifier sans ambiguïté certaines pratiques similaires pour les six sites, et d'expliquer par des causes communes ces pratiques qui s'avèrent avoir des visées promotionnelles plutôt que véritablement journalistiques.

Notre enquête a conduit en particulier à deux constats significatifs sur lesquels nous nous concentrerons ici. Premièrement, plus de la moitié des titres étaient écrits sous la forme « Mot-clé :... ». Deuxièmement, environ 90% des liens hypertextes observés pointaient vers des pages du même site et plus de la moitié avaient été générés automatiquement vers des pages dont la valeur ajoutée informationnelle était faible. Ces deux observations n'ont pas manqué de nous étonner puisqu'il n'y avait *a priori* aucune raison pour que les titres des articles soient construits de la même façon par les différents journalistes et puisque la vocation des liens était selon nous de pointer vers des contenus à forte valeur ajoutée. C'est pourquoi nous avons voulu nous pencher sur ces constats en particulier et trouver des éléments d'explication qui nous permettraient de répondre, au moins en partie, à la question qu'ils nous posaient : Pourquoi ces sites qui ont autant de visiteurs jouent-ils si peu le jeu du web en ne se saisissant pas de la zone de titre comme un moyen de se différencier et du lien hypertexte comme un moyen de s'ouvrir ?

Mot-clé, deux points

Le titre d'un article écrit par un webjournaliste peut remplir deux fonctions. D'une part, il constitue une information en lui-même : il dit quelque chose à propos de l'actualité. D'autre part, il incite le lecteur à vouloir en apprendre davantage. Il a donc un rôle informationnel et/ou un rôle de promotion. Il dit et il promet, il informe, incite, il prévient et suggère. D'une certaine façon, ces deux rôles s'opposent. Le premier implique de dire l'essentiel, le second de le suggérer. Le premier s'attache au fond, le second à la forme. Le premier consiste à diffuser une information jusqu'aux lecteurs, le second à attirer les lecteurs jusqu'à l'information. Une tension s'en suit nécessairement : à chaque fois qu'il rédige un titre, le journaliste doit se positionner dans le spectre qui va de l'un à l'autre de ces rôles. Au moment d'écrire le titre de son article, il lui faut arbitrer entre information et incitation : donner de l'information sans donner toute l'information et attirer le lecteur sans faire de fausses promesses.

Il existe sur le web de multiples interfaces où le titre d'un article, et parfois les premiers mots du texte, sont repris automatiquement. C'est ce que nous avons appelé dans un travail précédent le phénomène de « Multi-Une » (Sire, 2013). Sur Facebook, Google, Twitter, les agrégateurs de flux RSS, le titre de l'article et quelques mots apparaissent, parfois accompagnés d'une image ; le double rôle du titre est alors crucial.

Considérons d'abord le rôle informationnel. Une étude réalisée en 2010 par la compagnie Outsell a montré que 44% des utilisateurs américains de Google Actualités lisaient les titres et les accroches affichées sur l'interface sans cliquer sur les liens (Wauters, 2010). Dans une telle situation, si on considère que le but du journaliste est avant tout d'informer les citoyens à propos de l'actualité, on comprend qu'il est important que le titre soit clair, explicite, et qu'il contienne l'essentiel de l'information. Voyons maintenant le rôle promotionnel. La vocation des acteurs comme Google, Facebook, Twitter, etc. — tous ceux-là que Franck Rebillard et Nikos Smyrniotis nomment des « infomédiaires » (Rebillard et Smyrniotis, 2010) — est de se positionner sur le web comme des carrefours d'audience, chargés de mettre de l'ordre dans le chaos informationnel et de guider les internautes. Ils constituent des leviers considérables pour les éditeurs qui désireraient augmenter le trafic sur leurs pages. Ils braquent leurs projecteurs sur certains contenus qui deviennent de fait « plus publics » que les autres (Cardon, 2010). C'est pourquoi de nombreux éditeurs cherchent d'une part à ce que les titres de leurs articles apparaissent sur ces interfaces et, d'autre part, à ce qu'ils provoquent la venue d'un maximum d'internautes. Ce dernier point est une condition *sine qua none* de la bonne marche du modèle économique de nombreux éditeurs de presse en ligne qui comptent sur l'affichage d'encarts publicitaires, et la facturation de ces encarts au coût par affichage pour monnayer leur activité (Ouakrat *et al.*, 2010). Ceux là ont donc besoin d'avoir le plus grand nombre de visiteurs possible sur leurs pages. Ils cherchent à apparaître sur les interfaces des infomédiaires et à « susciter le désir d'expérience » (Sonnac, 2009), car si le lecteur se contente de lire le titre qu'il voit apparaître sur Google ou sur Facebook, alors l'éditeur ne pourra même pas monnayer son attention auprès des publicitaires.

Sachant cela, le journaliste est confronté à un dilemme. Soit il décide de dire l'essentiel de l'information dans son titre et privilégie la dimension informationnelle (exemple : *Goncourt : Houellebecq lauréat*). Soit il décide d'utiliser le titre comme un message adressé à l'internaute pour l'inciter à se rendre sur la page où il trouvera la totalité de l'information et où sa présence sera monnayée (exemple : *Un Goncourt sans surprise*). Au passage, il est important de noter que ce n'est pas parce que le journaliste privilégierait le rôle promotionnel du titre qu'il ne serait motivé que par des raisons économiques : certaines informations sont trop complexes pour être résumées par un titre, et doivent faire l'objet de développements plus longs, ce qui pourrait expliquer pourquoi le journaliste préfère utiliser le titre comme « appât ». Quoi qu'il en soit, au moment d'écrire son titre, le journaliste doit arbitrer entre les deux fonctions ou bien, sinon, réussir à remplir les deux fonctions à la fois, ce qui n'est pas impossible (il donnera alors l'information sans donner trop d'information, et incitera le lecteur à en savoir davantage sans pour autant ne rien lui dire au sujet de ce qu'il y a à savoir). Cela ne dépend bien sûr pas du seul journaliste. Un secrétaire de rédaction peut avoir la main sur la titraille, ainsi que ce qu'on appelle en général dans les rédactions du web le « *Front Page Editor* », en charge de la page d'accueil du site et, plus généralement, de tout ce qui apparaît sur la Multi-Une : les titres, les chapeaux, les premiers mots des articles. Le rédacteur en chef et les rédacteurs en chef adjoint peuvent eux aussi corriger un titre avant de valider l'article. La forme des titres dépend également de la stratégie éditoriale définie en amont et des consignes qui ont été données aux journalistes concernant la ligne éditoriale. Le journaliste n'est donc pas simplement confronté à un dilemme entre d'une part des impératifs économiques et d'autre part la volonté d'informer le plus grand nombre

d'internautes, mais également à des acteurs qui n'auront pas forcément le même avis que lui quant à la manière dont il faudrait composer avec ce dilemme. Il y aura forcément des jeux de pouvoir, des négociations, des compromis.

La tension est accrue par le fait que les infomédiaires les plus prescripteurs d'audience (Google fournit environ 50% de leur trafic à tous les sites de notre échantillon) sont algorithmiques. Ils hiérarchisent l'information de façon automatique en fonction de critères et de pondérations définis au préalable. Ces machines ne comprennent pas les jeux de mots et restent insensibles aux sirènes des titres qui seraient uniquement incitatifs. Elles « aiment » ce qui est parfaitement informatif. Il s'agit donc, pour traverser leur prisme, de rédiger un titre qui dira sans ambiguïté de quoi il est question. Il convient toutefois de considérer la problématique suivante : une fois que le titre aura traversé le prisme, il sera montré à l'internaute sur l'interface de l'infomédiaire, et alors l'internaute, qui n'est pas un algorithme, devra choisir s'il clique ou non dessus. Il faudrait donc l'inciter à cliquer, à défaut de quoi l'infomédiaire ne le reconduira pas vers la page du site de presse en ligne, aucune publicité ne sera affichée, et le journaliste aura donné à l'internaute une information sans pouvoir monnayer l'attention que pourtant l'internaute y aura bel et bien accordée. Autrement dit, il s'agit d'être assez informatif pour franchir le prisme des infomédiaires algorithmiques, et assez incitatif pour que cela puisse avoir un intérêt du point de vue économique.

Puisque c'est le plus gros pourvoyeur de trafic, gardons l'exemple de Google. Le titre idéal pour ce moteur de recherche qui est en position de monopole en France avec plus de 93% des parts de marché (AT Internet, 2014) comporte deux points précédés d'un mot-clé. Le mot-clé en question doit à la fois décrire sans ambiguïté le sujet dont il est question et être précisément le mot qui sera utilisé dans leurs requêtes par tous les internautes qui utiliseront le moteur de recherche généraliste Google Search ou le dispositif dédié Google Actualités pour se renseigner à propos de l'actualité dont il est question. Google met à disposition des éditeurs un outil nommé « Tendances de recherche » grâce auquel ils peuvent se renseigner à propos des mots-clés qu'il est préférable d'utiliser dans leurs titres. Le journaliste y a accès gratuitement (<http://www.google.com/trends/>). Il peut comparer plusieurs termes entre eux. Par exemple, il peut comparer « DSK » et « Strauss-Kahn » pour voir lequel de ces deux mots-clés est le plus utilisé sur le moteur Google, et employer celui-là, avant les deux points, dans son titre, pour maximiser l'apport de trafic en provenance de Google.

Le risque de cette logique est l'uniformisation des titres. Imaginons un instant que Google soit le seul pourvoyeur de trafic. Il serait fort probable dans ce cas que tous les journalistes consacrant un article à Dominique Strauss-Kahn et à l'affaire du Carlton choisissent de le nommer « *DSK : l'affaire du Carlton* », quelles que soient les différences qu'il y aurait entre leurs articles. Le lecteur ne disposerait alors plus d'aucune indication lui permettant de dissocier les angles de traitement choisis. Pourtant certains articles seront des articles d'analyse, d'autres de simples dépêches AFP, d'autres des interviews, d'autres des chroniques d'opinion, etc. Le titre contiendra certes l'essentiel de l'information concernant l'actualité, mais il n'informerait plus du tout à propos de l'article ; il aura perdu son rôle méta-informatif.

Une des autres limites de cette stratégie du titre « Mot-clé :... » est liée à l'information elle-même, qui risque d'être faussée précisément parce que le journaliste aura essayé de « plaire à l'algorithme » davantage qu'il n'aura tenté d'informer ou d'intéresser le lecteur. Dans un article du *Poynter*, la journaliste Kelly McBride a révélé quelles pouvaient être les dérives de l'influence de Google sur le choix des mots dans les titres (McBride, 2011). Elle a montré comment le traitement journalistique du projet de construction d'un centre musulman à deux pâtés de maison et demi du *Ground Zero* à New York avait pu être erroné par l'usage massif de l'outil « Tendances de Recherche ». Puisque les internautes formulaient la requête « *ground zero mosque* » sur le moteur, de nombreux journalistes utilisaient ce terme à des fins

de captation de trafic. Mais l'information donnée par les titres sous-entendait qu'une mosquée avait effectivement été construite sur le *Ground Zero*, ce qui était faux puisque la mosquée devait être construite beaucoup plus loin (les deux pâtés de maison étant de tailles considérables) et que le projet, porté par des musulmans modérés, n'était pas seulement de construire une mosquée, mais également un centre ouvert à tous comprenant une piscine, un terrain de basket, une salle de gym, un auditorium de 500 places, une librairie, un studio d'art, un restaurant et une école de cuisine séparés du lieu de culte. Ce point indique la limite de la stratégie qui consiste à reprendre dans le titre d'un article les termes utilisés par les internautes lors de leurs requêtes sur un moteur de recherche, et qui est pourtant plébiscitée par certains journalistes (Richmond, 2008 ; Niles, 2010). En effet, l'objectif de la requête effectuée sur un infomédiaire comme Google est d'exprimer un besoin d'information, tandis que le but d'un article est de traiter l'information. Le premier est une question, le deuxième une réponse. Dès lors, la stratégie qui consiste à reprendre dans la réponse les termes utilisés dans la question, parfaitement valable du point de vue de la promotion des contenus sur le web, peut conduire à certaines erreurs inacceptables du point de vue informationnel.

Notre observation en ligne montre que les titres sous la forme « Mot-clé :... » étaient surreprésentés en 2012 (56%), et l'étaient toujours, et de façon quasiment identique, en 2015 (55%). Plusieurs journalistes, interrogés en 2012 dans le cadre de notre thèse (Sire, 2015), nous ont expliqué qu'ils n'appréciaient pas cette forme de titre, et ont expressément regretté l'uniformisation qui en découle. Ils écrivent plus de la moitié de leurs titres sous la forme « Mot-clé :... » pour franchir le prisme des infomédiaires algorithmiques, et pour Google surtout, mais ils pensent que du point de vue strictement journalistique cela dessert le lecteur. Ils le font parce que cela leur a été, non pas ordonné, mais fortement suggéré, et parce qu'ils sont conscients qu'il est nécessaire d'atteindre un certain nombre de lecteurs pour que l'entreprise pour laquelle ils travaillent puissent continuer à exister.

Optimisation des titres

	Titre « Mot-clé :... » 2012	Titre « Mot-clé :... » 2015
Lexpress.fr	69%	52%
Lefigaro.fr	36%	40%
Lemonde.fr	40%	33%
Nouvelobs.com	57%	52%
Leparisien.fr	79%	76%
Lepoint.fr	55%	76%
MOYENNE	56%	55%

Les deux sites dont le trafic et la part d'accès directs sont les plus forts, *Le Figaro* et *Le Monde*, ont sensiblement moins de titres optimisés que les autres, en 2012 comme en 2015, tout en ayant plus d'un tiers de leurs titres optimisés. Les autres ont tous plus de la moitié de leurs titres qui sont optimisés, et même les trois quarts pour *Leparisien.fr* en 2012 et en 2015 et pour *Lepoint.fr* en 2015. Les pratiques des rédacteurs de ces six sites vont donc dans le sens contraire de l'idéal de diversité et de pluralisme qui a pu être mis en avant aux débuts du web. Sur le web, ce qui est visible, c'est ce qui est écrit sous cette forme « Mot-clé :... » et, donc, ce qu'on voit ici et là se ressemble. Cela ne veut pas dire qu'il n'y a pas d'autres sites sur lesquels les titres sont plus originaux, mais cela veut dire qu'on les voit moins, qu'ils sont moins visibles, et qu'il faut par conséquent fouiller sur le web pour trouver une certaine diversité dans la forme des titres. Cette observation rejoint les conclusions de différentes

études déjà effectuées à propos du pluralisme de l'information sur le web (Rebillard, 2006 ; Marty *et al.*, 2012)

Le constat que nous faisons est d'autant plus étonnant qu'il existe un moyen assez simple à mettre en œuvre techniquement pour dédoubler le titre. Le journaliste écrira dans ce cas deux titres : un titre pour les lecteurs, qui apparaîtra effectivement sur la page du site de presse en ligne, et un titre pour les infomédiaires algorithmiques, qui apparaîtra seulement sur l'interface des infomédiaires. C'est ce qui a été mis en place par exemple sur le site *BBC News* (Dick, 2011, p. 468) et sur le site *Lexpress.fr* (Sire, 2015). Ce double titre permet de reprendre la main sur le titre éditorial, présent sur la page, en revanche il ne résout pas la tension entre caractères informatif et incitatif, puisque le problème reste le même pour le deuxième titre, celui qui ne figure pas sur la page mais qui sera repris sur les infomédiaires : il devra être assez informatif pour franchir le filtre des algorithmes et assez incitatif une fois le prisme franchi.

Quelles que soient les raisons pour lesquelles la technique du double-titre est si peu mise en œuvre, il est clair qu'il existe une forte tendance à l'uniformisation des titres sur les sites de presse en ligne observés, majoritairement rédigés d'une façon optimisée pour les infomédiaires algorithmiques : « Mot-clé :... ». Ces titres sont informatifs, mais surtout, ils sont *fonctionnels*, dès lors que leur objectif est avant tout de « fonctionner » en servant de passeport permettant d'être « validés » par les algorithmes.

De soi jusqu'à soi-même,

Les liens hypertextes sont au cœur du projet du web, ils y servent à la fois de panneaux d'indication (le mot ou le groupe de mots sur le(s)quel(s) les liens hypertextes sont paramétrés, appelés « ancre », donnent des indications à propos de la destination du lien) et de portes, de passages (un clic sur le lien envoie effectivement vers cette destination). Ils sont la matérialisation du rêve de Vannevar Bush, qui, en 1945, dans son célèbre article *As we may think*, imaginait de relier les documents d'une même collection entre eux comme les idées et les souvenirs sont reliés dans le cerveau, en créant des ponts au sein même d'un document, sur une page en particulier, vers un autre document, une autre page, et non pas simplement par le biais des bibliographies. Dès le tout début des années 90, alors que le web n'en était qu'à ses prémices, Pierre Lévy, quant à lui, écrivait : « Dans les hypertextes, tout fonctionne à la proximité, au voisinage. Le cours des phénomènes y est affaire de topologie, de chemins. Il n'y a pas d'espace universel homogène où les forces de liaison et de déliaison, où les messages pourraient circuler librement. Tout ce qui se déplace doit emprunter le réseau hypertextuel tel qu'il est, ou est obligé de le modifier. Le réseau n'est pas dans l'espace, il *est* l'espace. » (Lévy, 1991, p. 61-69)

Pour le journalisme, l'usage du lien hypertexte a tout de suite été vu comme une opportunité. Cela permettrait notamment aux journalistes de citer leurs sources en renvoyant l'internaute qui le souhaitait vers la source en question. C'était également l'occasion de ne pas sans cesse répéter les faits déjà exposés et de se consacrer aux rebondissements les plus récents en renvoyant l'internaute qui le souhaitait aux articles déjà publiés, grâce auxquels il pourrait retracer la chronologie de tel ou tel sujet. En outre, les liens permettaient de diriger l'internaute vers certains documents et certains articles de fond écrits par des collègues, d'autres journalistes, travaillant pour certains d'entre eux pour d'autres sites de presse, où le lecteur trouverait de plus amples informations ou d'autres points de vue, d'autres angles, d'autres approches, et pourrait ainsi compléter la connaissance qu'il avait du sujet et se faire éventuellement son propre avis. Les chercheurs ont participé à cet engouement en expliquant la chance que constituaient les liens hypertextes pour les pratiques journalistiques. Selon Coddington par exemple, les liens auraient dû être utilisés dans les sites de presse en ligne

« comme un outil de mise en contexte et de référence, en pointant principalement vers des contenus intemporels, qui rassemblent des informations factuelles sur un sujet et font autorité » (Coddington 2012 *cit. in.* De Maeyer, 2013 p. 16). Cependant, il se trouve que vingt ans après les débuts du web, les liens, qui sont pourtant au cœur du projet du web, et dont les journalistes connaissent parfaitement les vertus (en 2015 il n'est plus possible de feindre de ne pas savoir créer un lien hypertexte ou de ne pas savoir à quoi cela peut servir quand on est webjournaliste), sont essentiellement utilisés par les sites de presse en ligne pour pointer vers eux-mêmes plutôt que vers l'extérieur, et pour être visibles sur Google plutôt que pour compléter et délinéariser le récit des événements d'actualité.

Juliette de Maeyer a consacré de nombreuses publications à l'usage effectué par les journalistes des liens hypertextes et aux discours tenus par eux à ce sujet (De Mayer, 2009, 2010, 2012, 2013). De son état de l'art extrêmement détaillé, il ressort que les liens sont finalement assez peu utilisés par les journalistes. Plusieurs études empiriques en attestent, qui concluent toutes que les liens sont peu employés par les journalistes pour citer leurs sources ou envoyer l'internaute vers des documents où il trouvera de plus amples informations (Tankard et Ban, 1998 ; Kenney *et al.*, 2000 ; Dimitrova *et al.*, 2003 ; Dimitrova et Neznanski, 2006 ; Stray, 2010).

Notre propre étude va également dans ce sens, en faisant apparaître que peu de liens pointent vers l'extérieur. De plus, notre observation a révélé que de nombreux liens n'étaient même pas édités par les journalistes, mais générés de manière automatique. Ces liens, que nous nommerons ci-après « liens automatiques », pointent tous vers des pages d'atterrissage. Celles-ci sont construites autour d'un mot-clé ou d'une thématique — un nom de lieu (Florange), un nom de personnalité (François Hollande), un nom d'événement (Roland Garros). En se rendant sur une page d'atterrissage, l'internaute voit apparaître les références des articles liés au thème auquel la page est consacrée (date, nom de l'auteur, lien, illustration) classées par ordre ante-chronologique. Ces pages sont conçues à la fois pour répertorier les archives à propos d'un même sujet et positionner la page sur le moteur de recherche Google. L'algorithme de ce dernier considère que chaque lien pointant vers une page est un vote pour cette page. Plus une page reçoit de votes, plus elle est considérée comme pertinente, et plus son vote, lorsqu'elle pointe elle-même vers d'autres pages, est significatif (Rieder, 2012 ; Cardon, 2013). Ainsi, certains éditeurs de contenus automatisent la création de liens sur des « pages filles » vers une « page mère » dans le seul but de positionner cette « page mère » en haut des résultats du moteur Google. C'est à cause de cette technique, dite *Page Rank Sculpting*, que sont créées ces pages d'atterrissage par les sites de presse en ligne. On y trouve répertoriés sous forme de liste de liens hypertextes (en général titre + accroche + image) l'ensemble des articles concernant un même sujet, classés dans l'ordre *ante-chronologique*. Dès que le nom « François Hollande » est utilisé par un journaliste dans un article, un lien est automatiquement créé vers la page d'atterrissage consacrée au Président de la République, ce qui permet de distribuer systématiquement l'autorité à cette page et de la faire remonter dans Google, où elle recevra potentiellement un trafic substantiel étant donné le nombre d'internautes qui effectuent la requête « François Hollande » sur Google. Une telle stratégie de *PageRank Sculpting* explique donc en grande partie l'« attitude protectionniste » des sites de presse en ligne (Steensen, 2011, p. 315).

Juliette de Maeyer ne fait pas le même constat, puisque d'après elle le faible nombre de liens vers l'extérieur relèverait d'une composante culturelle chez les journalistes plutôt que véritablement des stratégies de *PageRank Sculpting* : « Le lien semble dès lors représenter une matérialisation trop incarnée du mimétisme et de la circulation de l'information. À première vue, les réticences sont d'ordre économique : c'est d'ailleurs de cette manière que les journalistes les plus ouvertement opposés aux liens vers des concurrents le formulent — on ne va quand même pas leur envoyer du trafic, leur faire gagner des lecteurs. Mais à ce

calcul simple s'ajoutent des couches d'enjeux. Ainsi, le poids d'infomédiaires (Rebillard et Smyrnaio, 2010) comme Google et le mystère autour de leur mode opératoire exact créent des injonctions contradictoires : au sein d'une même rédaction, on voit certains affirmer que plus de liens externes, y compris vers des concurrents, sont une stratégie gagnante en termes d'optimisation pour les moteurs de recherche ; et d'autres défendent tout le contraire. S'il y a un effet économique des liens externes vers des concurrents, l'ampleur de celui-ci ou même son sens (s'agit-il d'un effet positif ou négatif ?) ne sont pas connus — en tout cas pas des journalistes — et, dans les deux rédactions étudiées, il ne semble pas exister de pressions directes de l'entreprise sur les journalistes pour produire ou ne pas produire de liens. Les réticences à ouvrir des routes directes vers la concurrence seraient donc plutôt d'ordre culturel, ou le fruit d'une culture d'entreprise bien intégrée par les journalistes. » (De Mayer, 2013, p. 248)

Les propos de Juliette de Maeyer vont clairement dans le sens de ce que nous ont dit certains experts du référencement sur Google qui accusent les journalistes d'être la cause de l'absence de liens externes (Sire, 2014). Cependant, les propos de Juliette de Maeyer ne concernent que les liens effectués par les journalistes directement. Pour ce qui est des liens effectués automatiquement vers des pages d'atterrissage, il nous semble évident qu'ils sont utilisés comme un outil de valorisation destiné au seul moteur de recherche.

Liens entrants et sortants

	Nombre de liens dans le corps de l'article		Nombre liens vers une page du même site		Nombre de liens vers une page d'atterrissage		Nombre de liens vers un autre site	
	2012	2015	2012	2015	2012	2015	2012	2015
Lexpress.fr	315	249	284 (90%)	196 (78%)	131 (42%)	37 (15%)	31 (10%)	53 (22%)
Lefigaro.fr	155	113	133 (86%)	90 (80%)	54 (35%)	38 (34%)	22 (14%)	23 (20%)
Lemonde.fr	1258	708	1202 (96%)	698 (99%)	1141 (91%)	576 (81%)	56 (4%)	10 (1%)
Nouvelobs.com	128	251	114 (89%)	171 (68%)	84 (66%)	92 (37%)	14 (11%)	80 (32%)
Leparisien.fr	188	226	177 (94%)	197 (87%)	74 (40%)	98 (43%)	11 (6%)	29 (13%)
Lepoint.fr	73	153	70 (96%)	146 (95%)	54 (74%)	114 (74%)	3 (4%)	7 (5%)
TOTAL	2117	1700	1980 (94%)	1498 (88%)	1538 (73%)	955 (56%)	137 (6%)	215 (12%)

Tableau 12. Moyennes du nombre de liens entrants et sortants

	Moyenne du nombre de liens par article		Moyenne du nombre de liens internes par article		Moyenne du nombre de liens sortant par article	
	2012	2015	2012	2015	2012	2015
Lexpress.fr	7,5	5,9	6,8	4,7	0,7	1,3
Lefigaro.fr	3,7	2,7	3,2	2,1	0,5	0,5
Lemonde.fr	30	16,9	29	16,6	1,1	0,2
Nouvelobs.com	3	6	2,7	4	0,3	1,7
Leparisien.fr	4,5	5,4	4,2	4,7	0,5	0,7
Lepoint.fr	1,7	3,6	1,6	3,5	0,07	0,1

Dans notre étude, 94% des liens en 2012 et 88% en 2015 pointent vers une page du site sur lequel ils se trouvent. Même si notre observation ne concerne « que » 504 articles, la tendance à l'égocentrisme hypertextuel apparaît clairement. Sur ces liens, 73% en 2012 et 56% en 2015 sont des liens qui n'ont pas été effectués par des journalistes mais générés automatiquement, et n'ont aucune vocation à caractère journalistique. Ils poursuivent essentiellement un rôle de promotion des pages d'atterrissage, et non des articles, dans les résultats des moteurs de recherche. Tandis qu'on observe des comportements extrêmement stables (la proportion de liens automatiques est de 74% pour Lepoint.fr en 2012 et 2015, et de 34% en 2012 et 35% en 2015 pour Lefigaro.fr), il semblerait que certains sites aient changé de politique, notamment Lexpress.fr qui est passé de 42% de liens automatiques en 2012 à seulement 15% en 2015. Renseignements pris, il s'avère que l'employé en charge d'optimiser le référencement des articles de Lexpress.fr sur Google a changé entre 2012 et 2015, ce qui pourrait expliquer cette différence.

La proportion de liens automatiques, même si elle reste forte, a sensiblement baissé entre 2012 et 2015, passant de 73% à 56%. Cependant, le nombre moyen de liens sortants par article est resté proche de 0,1, alors que le nombre moyen de liens vers soi-même dans chaque article a sensiblement augmenté en passant de 1,6 à 3,5. Cela montre bien que l'égocentrisme hypertextuel reste prégnant dans ces six sites d'information : la majorité des liens sont des liens générés automatiquement vers des passage d'atterrissage, et la majorité des liens qui ne sont pas générés automatiquement sont dirigés par les journalistes vers les pages de leurs propres sites.

Il semblerait que nous retrouvions ici une situation semblable à celle observée par Juliette De Maeyer sur le site belge d'information *7sur7* : « l'automatisation de la création de "faux liens externes" représente une intervention forte du marketing au sein des contenus éditoriaux, complètement intégrée dans le dispositif technique, sur laquelle les rédacteurs n'ont aucune prise. Avoir choisi de développer de telles fonctionnalités gonfle de fait le volume de liens de cette nature » (De Maeyer, 2013, p. 234). La chercheuse souligne que pourtant l'effort d'automatisation pourrait être fait en faveur de liens contextuels, qui approfondiraient l'information, mais que ce n'est pas le cas, ce qui montre que le caractère promotionnel des liens est privilégié à leur caractère informationnel.

Seule une enquête qualitative nous permettrait d'aller plus loin concernant l'usage des liens hypertextes. Il faudrait tâcher notamment de comprendre pourquoi les journalistes des

sites de notre panel utilisent si peu les liens pointant vers l'extérieur. Une politique éditoriale concernant les liens a-t-elle été définie en interne, ou bien les journalistes font-ils ce qu'ils veulent ? Est-ce faute de temps ? Est-ce faute d'intérêt ? Toutes ces questions nécessitent une enquête approfondie sur les causes du constat que nous faisons ici. Mais celui-ci, au moins, a le mérite d'être clair : les sites de presse en ligne adossés à des titres de presse imprimée payants d'information politique et générale ont tendance à pointer vers eux-mêmes, et la majorité des liens effectués vers eux-mêmes sont générés automatiquement ; ces liens ont des visées promotionnelles et non informationnelles ; ils visent à positionner des pages d'atterrissage sur Google plutôt qu'à donner une information à l'internaute qui est en train de lire l'article dans le corps duquel le lien se trouve.

Conclusion : les plus visibles sur le web le sont précisément parce qu'ils ne jouent pas le jeu du web

Voilà plus de dix ans maintenant que la presse en ligne a atteint ce que Pablo Boczkowski nomme la « phase de rationalité » (Boczkowski, 2010, p. 48), c'est-à-dire le moment où il ne s'agit plus de montrer qu'on est présent sur le web pour « être dans le coup » mais d'agir de manière à ce que cela soit profitable. Cette exigence de rentabilité génère inévitablement certaines pressions sur les journalistes, d'autant qu'ils se trouvent dans un univers extrêmement concurrentiel. Ils sont contraints « d'osciller entre logique industrielle et marchande d'un côté et logique citoyenne de l'autre ; [...] écartelés entre l'obligation de respecter des impératifs de vente, d'audience, et leur souci d'assurer au mieux l'autonomie de leur pensée » (Rieffel, 2008, p.103).

Pour être rentable, un site de presse qui a fondé une partie de son modèle d'affaires sur la publicité facturée au coût par affichage doit avoir de nombreux visiteurs. Pour avoir de nombreux visiteurs, il lui faut franchir le prisme des infomédiaires algorithmiques. Et pour franchir ce prisme, il lui faut écrire des titres informatifs, clairs, sans ambiguïtés, et faire des liens hypertextes vers ses propres contenus. Rien n'oblige les entreprises de presse en ligne à le faire, mais elles le font, car autrement leurs contenus risqueraient de ne pas être visibles, ou beaucoup moins. C'est ainsi que l'on peut dire qu'il existe un « pouvoir normatif » des infomédiaires (Röhle, 2008). Quand Google préconise de faire telle ou telle action, que ce soit en communiquant directement ou bien simplement parce que les éditeurs auront compris comment fonctionne son algorithme et ce que cela implique de faire, de nombreux éditeurs ajustent leurs comportements. Cela crée nécessairement une uniformisation des pratiques et, donc, des contenus.

Est-ce que c'est aux journalistes d'adapter leurs pratiques aux algorithmes de Google et Facebook, ou est-ce au contraire à Google et Facebook d'adapter leurs algorithmes pour que les éditeurs de presse puissent faire ce qu'ils veulent ? Cette question n'a pas, et ne peut pas avoir, une réponse évidente. Les éditeurs ajusteront leurs comportements en fonction de ce que font les concepteurs des infomédiaires algorithmiques, qui eux-mêmes ajusteront leurs comportements en fonction de ce qu'ils auront observé chez les éditeurs et les internautes. Cela peut aboutir à ce que des pratiques qui sont pourtant au cœur du projet du web, comme l'édition de liens hypertextes, soient détournées ou esquivées.

Ce que nous avons observé ne s'applique cependant qu'aux sites de presse en ligne adossés à des titres imprimés d'informations politique et générale en France. Ce n'est peut-être et sans doute pas la même chose pour les titres de presse nés en ligne, pour les blogs et pour les journalistes-programmeurs qui, certainement, se saisissent des potentialités dont les médias traditionnels, étant données des réticences en interne et des traditions paralysantes, ne se saisissent pas, ou peu, ou moins bien. On peut en effet faire l'hypothèse que les blogueurs, les journalistes-programmeurs et de manière générale les rédacteurs des titres de

presse nés en ligne utilisent les liens hypertextes à des fins informationnelles et qu'ils sont moins enclins à uniformiser leurs titres. Il faudrait le vérifier en effectuant une nouvelle étude empirique. Mais quoi qu'il en soit, les chiffres publiés par l'OJD concernant la fréquentation des sites web font apparaître sans ambiguïté possible que ce sont pour l'instant les sites adossés à des médias traditionnels qui reçoivent le plus grand nombre de visiteurs. L'impression suscitée par une telle observation est paradoxale et pourrait s'exprimer comme suit : vingt après l'arrivée du web dans les rédactions, il semblerait que les sites de presse en ligne les plus visibles sur le web soient ceux qui jouent le moins le jeu du web.

Bibliographie

AT Internet (2014), *Baromètre des moteurs de recherche*.

Boczkowski P.J. (2010), « Ethnographie d'une rédaction en ligne Argentine. Les logiques contraires de la production de l'information chaude et froide », *Réseaux*, n°160-161, p. 43-78.

Cardon D. (2010), *La démocratie Internet. Promesses et limites*, Paris, Seuil, 102 p.

Cardon D. (2013), « Dans l'esprit du PageRank. Une enquête sur l'algorithme de Google », *Réseaux*, 1, n°177, p. 63-95.

Coddington M. (2012), « This Week in Review: A Limp Response to Plagiarism, and a Proposed Tax to Save Newspapers », *Nieman Journalism Lab*, disponible ici : <http://www.niemanlab.org/2012/09/this-week-in-review-a-limp-response-to-plagiarism-and-a-proposed-tax-to-save-newspapers/>.

De Maeyer J. (2009), « Journalism practices in an online environment. An investigation into the evolving specificities of newswork in Belgium, as challenged by technological changes », *Journalism in crisis conference*, University of Westminster.

De Maeyer J. (2010) « Être journaliste dans un environnement 2.0 », *Les Cahiers du Numérique*, 6(1), p. 157-177.

De Maeyer J. (2012) « Towards a hyperlinked society: A critical review of link studies », *New Media & Society*, vol. 15, n°5, p. 737-751.

De Maeyer J. (2013), *L'usage journalistique des liens hypertextes. Étude des représentations, contenus et pratiques à partir des sites d'information de la presse belge francophone*, Thèse de doctorat, Université Libre de Bruxelles, Faculté de Philosophie et de Lettres, 295 p.

Dick M. (2011), « Search engine optimisation in UK news production », *Journalism Practice*, vol. 5, n°4, p. 462-477.

Dimitrova D. V., Neznanski M. (2006) « Online Journalism and the War in Cyberspace: a comparison between U.S. and international newspapers », *Journal of Computer-mediated Communication*, 12(1), p. 248-263.

Dimitrova D. V., Connolly-Ahern C., Williams A. P., Kaid L. L., Reid A. (2003) « Hyperlinking as Gatekeeping: online newspaper coverage of the execution of an American terrorist », *Journalism Studies*, 4(3), p. 401-414.

- Domingo D.** (2006). *Inventing online journalism. Development of the Internet as a news medium in four Catalan online newsrooms*, Universita Rovira i Virgili, Tarragona, Catalonia.
- Kenney K., Gorelik A., Mwangi S.** (2000), « Interactive Features of Online Newspapers » *First Monday*, 5(1).
- Lévy P.** (1991). « L'hypertexte, instrument et métaphore de la communication », *Réseaux*, vol. 9, n°46-47, p. 59-68.
- McBride K.** (2011), « SEO Makes It Too Late for Truth for ‘Ground Zero Mosque’ », *Poynter*, disponible ici : <http://www.poynter.org/latest-news/making-sense-of-news/105201/seo-makes-it-too-late-for-truth-for-ground-zero-mosque/>
- Niles R.** (2010), « Student Journalists Need to Learn SEO More Than They Need AP Style », *Online Journalism Review*.
- Ouakrat A., Beuscart J.-S., Mellet K.** (2010), « Les régies publicitaires de la presse en ligne », *Réseaux*, n° 160-161, p. 135-161.
- Rebillard F.** (2006), « Du traitement de l'information à son retraitement », *Réseaux*, vol. 3, n°137, p. 29-68.
- Rebillard F.** (2007), *Le Web 2.0 en perspective: une analyse socio-économique de l'internet*, Paris, L'Harmattan, 162 p.
- Rebillard F., Smyrnaiois N.** (2010), « Les infomédiaires, au cœur de la filière de l'information en ligne », *Réseaux*, vol.2, n°160-161, p. 163-194.
- Rieder B.** (2012), « What is in PageRank? A Historical and Conceptual Investigation of a Recursive Status Index », *Computational Culture*, 28 septembre 2012.
- Rieffel R.** (2008), « La métamorphose de l'information : de sa production à sa coproduction », in : Greffe X., Sonnac N. (dir.), *Culture Web*, Paris, Dalloz.
- Richmond S.** (2008), « How SEO Is Changing Journalism », *British Journalism Review*, 19, p. 51-55.
- Röhle T.** (2009), « Dissecting the Gatekeepers. Relational Perspectives on the Power of Search Engines », in : Becker K, Felix S. (dir.) *Deep Search. The Politics of Search beyond Google*, Innsbruck, StudienVerlag, p. 117-132.
- Sire G.** (2013) – « La Multi-Une : comment les sites de presse en ligne s'adressent au lecteur potentiel », *Recherches en Communication*, n°39, p. 75-89.
- Sire G.** (2014), « Référenceur et référencement : cachez ces pratiques que je ne saurais voir », *Sur le journalisme*, vol.3, n°1, p. 70-83.
- Sire G.** (2015), *Google, la presse et les journalistes. Analyse interdisciplinaire d'une situation de coopération*, Bruxelles, Bruylant/Concurrences, coll. « Sciences Politiques ».
- Sonnac N.** (2009), « L'économie de la presse : vers un nouveau modèle d'affaires », *Les Cahiers du Journalisme*, 20, p. 22-42.

Steensen S. (2011) « Online Journalism and the Promises of New Technology » *Journalism Studies*, 12(3), p. 311-327.

Stray J. (2010) « Why Link Out? Four journalistic purposes of the noble Nieman Journalism Lab », *Nieman Journalism Lab*, <http://www.niemanlab.org/2010/06/why-link-out-four-journalistic-purposes-of-the-noble-hyperlink>.

Tankard J. W., Ban H. (1998), « Online Newspapers: living up to the potential? », *Annual conference of the AEJMC*, Baltimore, MD.

Wauters R. (2010), « Report: 44% of Google News Visitors Scan Headlines, Don't Click Through », *TechCrunch*, <http://techcrunch.com/2010/01/19/outsell-googlenews/>