Linx

Revue des linguistes de l'université Paris X Nanterre

52 | 2005

Lexique, terminologie, discours

Les mots et l'internet

Sandrine Reboul-Touré

Édition électronique

URL : <http://journals.openedition.org/linx/237>

DOI : 10.4000/linx.237

ISSN : 2118-9692

Éditeur

Presses universitaires de Paris Nanterre

Édition imprimée

Date de publication : 1 juin 2005

Pagination : 193-203

ISSN : 0246-8743

Référence électronique

Sandrine Reboul-Touré, « Les mots et l'internet », *Linx* [En ligne], 52 | 2005, mis en ligne le 31 janvier 2011, consulté le 30 avril 2019. URL : <http://journals.openedition.org/linx/237> ; DOI : 10.4000/linx.237

Département de Sciences du langage, Université Paris Ouest

Les mots et l'internet

Sandrine Reboul-Touré

SYLED – CEDISCOR, Université Paris 3 – Sorbonne nouvelle

La diffusion remarquable de l'internet dans notre société a causé des changements de différents ordres. Sur le plan linguistique, nous souhaitons montrer tout d'abord que ce moyen de communication a entraîné dans son sillage de nouvelles dénominations qui s'insèrent peu à peu dans le lexique usuel. Il est d'ailleurs possible aujourd'hui d'identifier « le vocabulaire de l'internet »¹. En comparant les textes issus des commissions de terminologie et de néologie et des textes circulant sur l'internet, nous soulignerons la présence notable de variantes. Par ailleurs, nous pouvons constater que le réseau internet donne au mot une place tout à fait privilégiée. Nous analyserons dans une deuxième partie les mots hors contexte lors de leur utilisation par la fonction « recherche » qui est fondamentale dans la collecte des informations². L'utilisateur doit en effet sélectionner les mots-clés qui lui paraissent pertinents pour obtenir l'information désirée. C'est donc un véritable index – une *table alphabétique*³ ouverte – qui est mis à la disposition de l'internaute grâce à cette fonction. Enfin, nous relèverons une autre mise en valeur du mot, en discours, notamment au travers du lien hypertexte. La couleur et/ou le soulignement pointe le mot et appelle un retour, sous la forme d'une modalisation autonymique. Le discours se dédouble alors sur plusieurs *couches* de textes.

¹ *Vocabulaire de l'internet*, J.O. du 20 mai 2005.

² Notre regard ne sera pas celui du spécialiste des réseaux.

³ *Petit Robert électronique*, 1996.

1. Le vocabulaire de l'internet

Parmi les textes officiels nous trouvons le *Vocabulaire de l'informatique et de l'internet* (mars 1999) ainsi que le *Vocabulaire de l'internet* (mai 2005) issus de la Commission générale de terminologie et de néologie :

En France les tâches d'aménagement de corpus ont été confiées aux Commissions ministérielles de terminologie (CMT), coordonnées par le Commissariat général de la langue française et plus récemment par la Délégation générale à la langue française. [...] Le but est d'éliminer les anglicismes employés dans le secteur de compétence d'un ministère donné [...] Parfois aussi on souhaite remplacer un terme qui prête à confusion ou qui est jugé mal formé. (Humbley, 2001, p. 127).

Nous pouvons ici identifier des « événements linguistiques » dans la mesure où il y a nomination de vocabulaire et que ce processus a lieu en amont, « alors que l'évènement discursif concerne l'approche configurationnelle de ce qui est dit dans les énoncés d'archive sous une forme attestée, l'évènement linguistique se définit en amont d'un tel sens advenu » (Guilhaumou, 2002). Les textes publiés au *Journal Officiel* présentent des termes et des définitions, ainsi que des tables d'équivalence anglais/français, français /anglais (Cf. des exemples en annexe). Cette source énonciative, clairement identifiée, propose des dénominations *a posteriori* qui entrent en concurrence avec des termes en usages.

1.1. Des variantes

L'identification d'un vocabulaire de l'internet permet de s'interroger sur le concept de dénomination. Un nom est donné à une réalité nouvelle par un acte de baptême du côté de la terminologie avec une volonté de fixation au niveau du lexique. Il existe cependant en amont des dénominations forgées par des concepteurs et ces dernières sont largement diffusées par différents supports médiatiques. Plusieurs mots coexistent donc pour un même référent. Or, parmi ces coréférents, il est difficile d'identifier une seule dénomination et des désignations : il y a plusieurs dénominations. La distinction établie par G. Petit entre « dénomination de droit » et « dénomination de fait » nous paraît particulièrement éclairante :

La pathologie que manifestent certains bovins ayant consommé des farines animales est exprimée régulièrement par *encéphalopathie spongiforme bovine* (siglée *ESB*) ou bien par *maladie de la vache folle* (ou *vache folle* tout court). Ces deux séquences établissent leur fonction coréférentielle avec la même récurrence et la même stabilité. Elles apparaissent dans des contextes discursifs différents, plutôt scientifiques pour la première et non scientifiques pour la seconde. On les rencontre toutefois en cooccurrence dans le discours médiatique destiné au grand public. Ayant un comportement sémiotique identique au regard de leur catégorie référentielle, elles répondent *a priori* aux critères de la dénomination.

Elles se distinguent toutefois du fait qu'*encéphalopathie spongiforme bovine* est validée terminologiquement, contrairement à *maladie de la vache folle* qui relève davantage de l'usage courant. Pour la première nous parlerons de dénomination de droit, et de dénomination de fait pour la seconde. La dénomination de droit est une relation référentielle institutionnellement reconnue, normalisée par une instance légitimante [...]

La dénomination de fait se développe dans l'usage en dehors des réseaux institutionnels et ne requiert d'autre légitimation que celle de la communauté des locuteurs. (Petit, 2001, 106-107)

Dans le vocabulaire retenu, *causette*, *frimousse*, *fouineur* sont proposés comme équivalents de *chat*, *smiley*, *hacker* (Cf. annexe) ; on note ici une répartition entre dénominations de droit et dénominations de fait. Les premières sont très souvent reprises comme issues de la terminologie et on les trouve plutôt dans un cadre définitionnel. Les secondes sont employées au fil des discours comme par exemple :

À l'occasion de notre série de chats sur la Constitution européenne, nous avons reçu jeudi après-midi Annick Le Petit, porte-parole du Parti Socialiste, favorable au oui (Libération, 4 novembre 2004)

Un « hacker » de 17 ans interpellé (*Le Figaro*, 11 juillet 2003)

Notons que *chat* est la dénomination retenue par certains quotidiens en ligne (*France-Soir*, *Le Monde*, *Libération*...). Parallèlement à l'internet, les dénominations autour du courrier électronique sont très nombreuses : *courrier électronique*, *mail*, *e-mail*, *mél.*, *courriel*. Les propos de la Commission, ci-dessous, soulignent à quel point la variation est – pour certains – difficilement acceptable :

Par cette publication [*Journal officiel* du 20 juin 2003.], la commission générale a souhaité mettre un terme à une période d'hésitation. Utilisé depuis un certain temps, notamment au Québec, ce néologisme [*courriel*] s'est progressivement répandu dans l'usage français pour désigner le courrier électronique, qu'il s'agisse, le plus souvent, du message lui-même, ou, par extension, de la messagerie utilisée. [...]

Évocatrice, avec une sonorité bien française, le mot *courriel* est largement utilisé dans la presse et concurrence avantageusement l'emprunt à l'anglais *mail*.

La commission générale se range donc à la proposition québécoise désormais consacrée par l'usage, tout en maintenant la forme *courrier électronique* comme synonyme. En effet, dès 1997, elle avait publié comme équivalent français de *e-mail*, le terme *courrier électronique*, qui reste parfaitement adapté.

Enfin, le symbole *Mél.*, qui n'a jamais été proposé comme terme équivalent de *e-mail*, reste utilisable, comme symbole et jamais comme nom, devant une adresse électronique, comme Tél. s'emploie devant un numéro de téléphone.

Délégation générale à la langue française et aux langues de France,
<http://www.culture.gouv.fr/culture/dglf/terminologie/courriel.htm>

Ces phénomènes invitent à s'interroger sur l'articulation entre la norme et l'usage : « qui est aujourd'hui le dépositaire de l'usage légitime, et qui a vocation à venir remplacer la Cour de l'ancien régime pour permettre [un] étalonnage des usages [?]. » (Kerbrat-Orecchioni, 2003). C'est pourquoi :

Quelques incorrigibles puristes mis à part, il est incontestable que l'on s'oriente aujourd'hui vers une conception assouplie, moins monolithique et plus relativiste de la norme ; une conception surtout *adaptative*, qui intègre pour de bon le fait que les usages varient et doivent varier en fonction des situations et des domaines. (*idem*)

De nouveaux éléments

La télématique a laissé place à l'internet et le paradigme de mots en *-tique* est tombé en désuétude. Aujourd'hui, il est possible d'isoler deux éléments qui sont à l'origine de paradigmes ouverts : *cyber-* et *e-*. Le mot *cybernétique* réapparaît au milieu du XX^e siècle pour désigner l'étude des processus de contrôle et de communication chez l'être vivant et la machine (d'après le mathématicien américain, N. Wiener, 1948)⁴. La troncation *cyber-* est à l'origine de créations lexicales dont certaines sont lexicalisées (*cyberespace*, *cybernaute*, *cybercafé*) mais *cyber-* ne conserve pas la valeur dénomminative de *cybernétique* et se charge d'autres valeurs sémantiques :

D'après l'anglo-américain, CYBER- est devenu (1993-94) un élément de mots composés à la mode, avec l'idée d'automatisme informatique, de robotique. Il s'applique surtout par calque de l'anglo-américain, mais aussi avec des formations spécifiques au français. Les composés anglo-américains s'appliquèrent d'abord aux automatismes : *cyborg*, de *cyber-* et *org(anism)*, signifie « organisme électronique humanoïde ». Puis, ils se spécialisent dans le domaine de la « toile » (le *web* ; Internet). EX. CYBERESPACE n.m., calque de *cyberspace* ; CYBERNAUTE n. empr. à l'américain, de *cyber-* et *(astro)naute*, pour personne qui se déplace [« surfe »] dans le cyberspace ». CYBERCAFÉ n.m., « café où des ordinateurs sont connectés au réseau Internet », est un composé français. La liste des composés est ouverte.⁵

Le paradigme est effectivement ouvert et nous trouvons des réalisations parfois inattendues : cyberacteur, cyberavocat, cyberbricoleur, cybercarte, cybercommerçants, cybercrime, cyberemploi, cybermarché, cyberpapy, cyberprof... L'emploi de *cyber-* est présenté comme « à la mode » or la mode est éphémère et il nous semble que le *e-* (troncation de *electronic*) puisse être à l'origine d'un nouveau paradigme ouvert. Les créations les plus souvent reprises sont le e-mail et le e-commerce (« Le palmarès 2003 du e-commerce français », titre dans *L'Expansion*, 18/12/2003). De nombreux sites utilisent l'élément *e-* pour leur adresses :

e-alsace, e-anglais, e-bahut (aide aux devoirs), e-campings, e-compil (pour la musique), e-corse, e-droit, e-fonctionnaires, e-gravures, e-historia, e-hotellerie, e-

⁴ Dictionnaire historique de la langue française.

⁵ *Idem*.

littérature, e-picardie, e-prénoms, e-santé, e-timbre, e-voyance... Ce paradigme semble bien résumé par le e-quelque chose proposé comme entrée dans le sabir cyber (Le Diberder, 2001). Remarquons que le e- est la troncation d'un emprunt car il n'est pas accentué ; il peut donc être prononcé [i] et donc parfois écrit *i* : *i-commerce*, *i-expo*, *i-mode*, *i-télé*, *i-ville*...

Néologie

Dans un domaine aussi foisonnant d'innovations, la création linguistique est dynamique. Parmi les créations récentes, notons *blog* emprunté à l'anglais. Le blog est un site internet, « nouvelle génération⁶, véritable caméléon, [qui] convient à tous : journal intime, site privé de la famille, site d'une communauté de passionnés, support de communication pour l'entreprise... »⁷. *Blog* est une troncation atypique de la composition *weblog* de *web* (*World Wide Web* ou toile d'araignée mondiale) et de *log* (journal de bord) :

Un weblog [...] est un site web proposant un journal en ligne tenu par une ou plusieurs personnes. [...] Les premiers weblogs sont apparus aux États-Unis à la fin des années 1990. Ils se composaient d'un carnet de bord recensant les hyperliens (vers des pages Web) que l'auteur avait jugé intéressants, accompagnés de commentaires concernant ces pages. Les premiers weblogs francophones apparaissent en 1996. [...] Par la suite le mot fut employé pour désigner la pratique du journal (plus ou moins) intime en ligne ou du carnet de notes publié électroniquement. Le point commun de toutes les formes de weblogs est qu'on y retrouve à intervalles irréguliers les impressions et sentiments de l'auteur du blog sur des sujets variés (sauf dans le cas des photoblogs...)⁸.

⁶ Dans une certaine mesure, nous sommes en train de vivre la "bloguisation du Web". Pas seulement [du fait] du contenu qu'on y trouve, mais de la forme que [le blog] y prend. Via RSS d'abord [RSS (Really Simple Syndication) est un moyen pour extraire d'un site web du contenu régulièrement mis à jour]. Le format RSS existait avant les blogs, et surtout avant qu'ils connaissent la popularité qu'ils ont aujourd'hui. Mais l'adoption massive par les blogueurs – et surtout par les outils de blogging – du format RSS (et/ou Atom) ont eu pour effet d'inciter – ou de contraindre – des sites qui ne sont pas des blogs à adopter ce format désormais presque universel.

Ce qui est vrai de RSS l'est également pour la manière dont sont présentés et publiés les blogs. Pour faire simple, indépendamment de son style ou de son contenu, un blog a trois caractéristiques principales :

- . les informations élémentaires y sont publiées de façon ante-chronologique (les plus récentes en haut)

- . elles sont horodatées et signées chacune par un auteur unique, en général clairement identifié (même s'il demeure anonyme)

- . elles sont complétées chacune d'un dispositif permettant au lecteur de les commenter

Le magazine du blogging, 5 mars 2004, http://www.pointblog.com/past/cat_analyses.htm

⁷ <http://www.u-blog.net/site/>

⁸ Wikipédia.

Du fait du développement notable de ce nouveau genre (?) de site, le néologisme *blog* est très utilisé par les internautes et on relève des dérivés et des composés : il existe des *blog(g)ueurs* ou *bloggers*, des *blogueuses*, des *journalistes-blogueurs* (par exemple Pierre Assouline avec le blog « La république des livres », *Le Monde*⁹) mais aussi la *blogosphère* parisienne, la *bloguisation*... Nous avons aussi trouvé un *lexicoblogue*¹⁰ qui propose d'éclairer les internautes sur les mots liés à l'univers des blogs. En France les Commissions¹¹ viennent récemment de proposer, *bloc-notes* :

bloc-notes, n.m.

Forme abrégée : bloc, n.m.

Domaine : Informatique/Internet.

Définition : Site sur la toile, souvent personnel, présentant en ordre chronologique de courts articles ou notes, généralement accompagnés de liens vers d'autres sites.

Note : La publication de ces notes est généralement facilitée par l'emploi d'un logiciel spécialisé qui met en forme le texte et les illustrations, construit des archives, offre des moyens de recherche et accueille les commentaires d'autres internautes.

Équivalent étranger : blog, web log, weblog.

(*Vocabulaire de l'internet*, J.O. du 20 mai 2005).

Notons la forme abrégée *bloc* [blok] très proche phonétiquement de [blog], *blog*...

2. Le mot : un sésame

La masse d'informations qui circule sur l'internet est une grande richesse mais il faut pouvoir y accéder. L'internaute ne sait d'ailleurs pas vraiment ce qu'il peut trouver sur la toile d'où des navigations parfois hasardeuses. Le problème de la recherche intelligente¹² est un point de réflexion largement débattu et ce n'est que très modestement que nous voudrions souligner ici l'importance donnée aux mots pour la recherche d'informations sur l'internet ainsi que la nécessaire réflexion autour des relations sémantiques qui structurent le lexique. Les concepteurs de recherche signalent d'ailleurs très clairement leurs préoccupations linguistiques :

⁹ <http://passouline.blog.lemonde.fr/>

¹⁰ *Blogable, Blogage, Blogeoisie, Blogobille, Blogoboule, Blogogeois/geoise, Blogogeoisie, Blogocentrique, Blogoliste, Blogoréalité, Blogorrbée, Blogoscopie, Blogosphérique, Blogostipation, Blogue mobile, Bloguitude, Bloguiversaire, Egoblogue, Métablogue, Moblogue, Photoblogue*, embruns.net/etude_du_blogue/000034.html

¹¹ Au Québec, « Le terme *blogue*, forme francisée de *blog*, a été proposé par l'Office québécois de la langue française (octobre 2000) sur le modèle de *bogue*, pour remplacer les termes anglais *blog* et *weblog*, très employés en français. Le mot *blogue* a permis la création de dérivés comme *bloguer* et *blogueur* qui sont de plus en plus répandus », Office québécois de la langue française, <http://www.olf.gouv.qc.ca/ressources/bibliotheque/dictionnaires/Internet/fiches/8370242.html>

¹² Samier, H., Sandoval, V. (1999), *La recherche intelligente sur l'internet et l'intranet*, Hermès sciences publications.

Le bien-fondé des méthodes de recherche textuelles est douteux : des pages qui intéresseraient beaucoup un utilisateur ne contiennent pas le mot qui a été transmis au moteur de recherche, tandis que d'autres pages qui contiennent le mot spécifié sont sans intérêt. Cette seconde difficulté provient de l'abondance de synonymes et de mots polysémiques [...] : une demande d'information sur les « automobiles » laissera de côté une kyrielle de pages qui évoquent les « voitures ». En raison des polysémies, une simple requête sur « jaguar » donnera des milliers de pages concernant, entre autres, l'automobile de ce nom, le félin et une équipe de football américain.

On améliore les recherches en utilisant les relations sémantiques entre les mots. Certains moteurs de recherche utilisent de telles relations, nommées « réseaux sémantiques », identifiés par les linguistes. À la réception de la requête « automobile », un moteur de recherche couplé à un réseau sémantique détermine d'abord qu'« automobile » est un mot équivalent à « voiture », puis il donne la liste des pages qui contiennent l'un des deux mots. Toutefois, cette méthode est à double tranchant : elle résout le problème des synonymes, mais aggrave celui de la polysémie.

De surcroît, l'établissement des réseaux sémantiques est une tâche redoutable car [...] les néologismes fleurissent. (*Recherche intelligente sur l'internet*).

Il est donc difficile pour les concepteurs de moteur de recherche d'identifier les sites les plus pertinents. Ils peuvent faire intervenir aussi les liens (« l'hypothèse [étant] que chaque lien matérialise une parenté sémantique entre la page initiale et la page vers laquelle il pointe », *idem*) ou encore les citations (la principale mesure de l'importance des articles [étant] le 'facteur d'impact', qui juge les publications d'après le nombre de références que les autres articles en font », *idem*).

Du côté de l'utilisateur, la recherche des mots-clés pose aussi des problèmes. En effet, on peut concevoir le mot-clé comme devant synthétiser les informations présentes dans un texte¹³. Or pour trouver un texte particulier sur l'internet mieux vaut isoler le mot le moins utilisé :

La situation est déstabilisante, puisqu'il faut présumer une chaîne de caractères qui pourrait apparaître dans un document. L'idéal à atteindre, c'est la chaîne de caractères suffisamment spécifique pour n'apparaître que dans le document souhaité.

En ce sens, le bon mot-clé est celui qui est suffisamment adéquat mais surtout suffisamment peu fréquent pour que le moteur de recherche l'ait indexé peu de fois. Peu de pages indexées avec ce mot-clé, mais des pages adéquates. [...]

¹³ « Une grande partie des apprentissages relatifs à la maîtrise des textes d'information suppose la maîtrise des mots-clés entendus comme des éléments qui intègrent des contenus de sens. De la sorte, il est possible de résumer, de structurer des textes, de synthétiser des connaissances. », Commission « français et informatique », (2000).

Il faudra absolument utiliser les termes hyponymes pour que la recherche soit rentable [...].

Or la recherche des hyponymes est directement déterminée par la compétence lexico-sémantique de l'élève. Plus son vocabulaire sera riche, plus il pourra présumer le terme adéquat qui permettra de détecter la page spécifique à son attente. Dans le cas où il ne connaît rien encore du domaine dans lequel il mène sa recherche, on se trouve là devant ce que les pédagogues nomment un paradoxe pragmatique¹⁴ : « Utilisez comme mot-clé un terme très spécifique que vous ne connaissez pas encore ».

(Berten, 2000).

La tension semble due à deux approches antagonistes, le repérage formel de caractères¹⁵ face aux choix sémantiques d'un individu. Les mots-clés sont utiles pour les outils de recherche sur l'internet qui se répartissent entre les moteurs de recherche et les annuaires. Cette différence correspond à la méthode de référencement des sites web :

Pour un moteur de recherche, il suffit de donner l'adresse du site web, qu'un logiciel (ou robot) vient visiter. Par rapport à son contenu, un score est automatiquement donné à ce site pour les différents mots-clés susceptibles d'être employés par l'internaute. L'annuaire s'appuie sur des hommes et sur leur jugement : le créateur du site donne l'adresse de celui-ci. Ensuite, quelqu'un vient visiter ce site web et décide ou non de l'inclure dans les résultats de l'annuaire¹⁶.

Dans ce dernier cas, les résultats sont plus pertinents. C'est sans doute la combinaison de plusieurs mots-clés qui permettra d'obtenir un meilleur résultat. La fonction « recherche » conduit à l'élaboration d'un vaste index fluctuant que chaque internaute module selon ses propres intérêts. Afin d'élargir le paradigme de mots-clés, un méta-moteur de recherche comme Kartoo¹⁷ est une aide précieuse. Par exemple, une recherche pour le mot *hyperonyme* donne les résultats suivants : *hyponyme* et *relation sémantique* (qui sont mis en relief), *exercices*, *antonyme*, *synonyme*, *méronymes*, *chat*, *chien*, *mammifère*, *animal*, *index*, *siège*, *lexicologie*, *terme*, *linguistique*, *base*, *dictionnaire*, *multilingue*. Notons aussi l'existence de moteur de recherche et de navigation sémantique comme

¹⁴ Remarquons que le paradoxe pragmatique se situe plutôt au niveau de l'énoncé : « Lorsque le contenu intrinsèque de l'énoncé se trouve [...] inadapté à ses conditions situationnelles d'utilisation, ou contredit par ce qu'implique son énonciation [...], F. Recanati cite « je ne sais pas écrire » énoncé par écrit [...], on parle alors de 'contradiction' ou de 'paradoxe pragmatique' ». Kerbrat-Orechioni, C. (1980) : *L'énonciation, De la subjectivité dans le langage*, Armand Colin, 1980, p. 187.

¹⁵ Nous avons constaté que si l'internaute entre un mot comportant une faute d'orthographe, sa recherche peut ne pas aboutir ou bien présenter tous les sites comportant la faute commise.

¹⁶ Dictionnaire, à l'aide.com, <http://www.alaide.com>

¹⁷ « Kartoo est un méta-moteur de recherche qui présente ses résultats sous forme de carte. Les sites trouvés sont représentés par des pages plus ou moins grosses en fonction de leur pertinence. Entre ces sites figurent des thèmes qu'il suffit de cliquer pour préciser [la] recherche, <http://www.kartoo.com/>

*Sinequa*¹⁸ qui fonctionne simultanément sur des stratégies de recherche utilisant la linguistique (lexicologie et sémantique), les statistiques, la mathématique et la recherche structurée. Les concepteurs exposent clairement leurs choix linguistiques et ce qu'ils entendent par :

- la recherche linguistique, avec l'analyse linguistique [qui] permet de retrouver les mots sous toutes leurs formes possibles (féminines, plurielles, conjuguées), et d'étendre la recherche aux mots synonymes ou de même racine étymologique. Cette recherche offre également des modules de recherche floue ou de phonétisation,
- la recherche lexicale, la recherche d'un mot ou d'un ensemble de mots, avec des opérateurs booléens (ET, OU, SAUF). A ces opérateurs classiques s'ajoutent des opérateurs d'adjacence, de proximité entre les mots, et une sensibilité paramétrable aux majuscules ou aux accents
- la recherche sémantique, une recherche [qui] peut s'effectuer à partir du sens de la requête, c'est-à-dire en recherchant les mots sémantiquement proches de ceux employés dans la requête. La recherche s'appuie alors sur un dictionnaire sémantique qui, à chaque mot de la langue, associe leurs différents sens. <http://www.sinequa.com/html-fr/fr-technologie.presentation.php>

Bien que nous soyons au cœur d'un multimédia, nous voyons ici toute l'importance accordée aux mots : ceux sont eux qui permettent l'accès à l'information.

3. Lien hypertextuel

La mise en valeur des mots sur l'internet prend aussi une autre forme, à l'intérieur des textes, avec la création possible de liens hypertextuels¹⁹. Les mots sont alors mis en évidence par une typographie particulière, la couleur – le plus souvent bleu – et/ou le soulignement et le gras. Dans le cadre du discours de transmission des connaissances, corpus privilégié dans notre recherche, les mots retenus comme liens sont les termes spécialisés. Ainsi, certains éléments d'une terminologie sont exhibés et explicités non plus à l'intérieur de la phrase²⁰ mais dans un ailleurs textuel. Comme par exemple dans les extraits d'un dossier scientifique du CNRS sur l'eau :

L'eau est un corps dont l'unité de base est une **molécule** : la molécule d'eau. La molécule d'eau est formée d'un **atome** d'oxygène relié à deux atomes d'hydrogène. On la note H₂O (H pour atome d'hydrogène et O pour atome d'oxygène). <http://www.cnrs.fr/cw/dossiers/doseau/decouv/proprie/1atomMole.html>

Lorsqu'on clique sur **molécule** apparaît un encart dans lequel le mot est défini ; il en est de même pour **atome** :

¹⁸ <http://www.sinequa.com/html-fr/fr-index.php>

¹⁹ Le lien hypertextuel permet, par un clic de souris, de passer rapidement à une autre partie du document ou bien à une autre page, sur le même site ou sur un autre site. On trouve les variantes suivantes : *lien hypertexte*, *lien*, *hyperlien*

²⁰ Voici quelques exemples « traditionnels » de vulgarisation : le terme est explicité au fil du discours. « La première méthode de transfert direct fut l'introduction mécanique d'ADN dans des protoplastes (cellules dont on a ôté la paroi pectocellulosique) ». *OGM et consommateurs*
« La possibilité de transfert croisé se pratique chez les plantes allogames c'est-à-dire issues de reproduction sexuée par des gamètes provenant de parents différents ». *Le mensuel de l'homéopathie*.

molécule : groupement stable constitué d'un nombre fini d'atomes liés entre eux et ayant une individualité propre.

atome : fragment microscopique de matière qu'il est encore possible de différencier de ses congénères même lorsqu'il est isolé ; à la base de presque toute la matière de l'univers ; un atome est constitué d'un noyau central autour duquel gravitent des électrons.

Cette dimension hypertextuelle nous invite à repenser certains aspects de l'écriture de la vulgarisation scientifique (Reboul-Touré, 2004). En effet, la reformulation, une des caractéristiques linguistiques fondamentales du discours de vulgarisation, est bouleversée dans sa présentation formelle puisqu'elle est, le plus souvent, décalée sur une autre page. Le lien hypertextuel nous apparaît comme une nouvelle marque de l'activité métalinguistique participant aux processus discursifs de la vulgarisation :

Les processus discursifs mis alors en jeu sont intéressants au double point de vue sémiotique et sémantique : en fonction de l'occurrence de termes scientifiques dans le discours de vulgarisation, les traces de l'activité métalinguistique repérables dans l'énoncé seront plus ou moins nombreuses et explicites, caractérisant l'aspect sémiotique de la vulgarisation. L'observation de la relation sémantique qui unit (en langue) les segments mis en relation de paraphrase dans le discours caractérise, de son côté, l'activité discursive de la vulgarisation... (Mortureux, 1982, p. 48)

Le lien, typographiquement identifiable, joue un rôle linguistique : il remplace des indices de reformulation comme *c'est-à-dire*, *ou*²¹. Il s'apparente alors à une ellipse. Il se présente aussi comme une marque invitant à la modalisation autonymique²² virtuellement réalisable : le mot utilisé comme lien est employé en usage dans la phrase (« la molécule d'eau est formée d'un **atome** d'oxygène... ») et le soulignement permet à l'internaute de revenir sur le terme pour obtenir un approfondissement et donc de percevoir le mot en mention (« **atome** : fragment microscopique de matière qu'il est encore possible de différencier de ses congénères même lorsqu'il est isolé ; à la base de presque toute la matière de l'univers ; un atome est constitué d'un noyau central autour duquel gravitent des électrons »).

L'effervescence technologique autour de l'internet rencontre un écho en linguistique. Tout d'abord la créativité lexicale est très dynamique, de nouveaux éléments sont forgés et le flot des néologismes est loin d'être tari. Des variantes coexistent car différentes sources énonciatives véhiculent des dénominations différentes : ce phénomène permet d'observer les tensions entre la norme – les normes ? – et les usages et de s'interroger sur les nouvelles instances normatives, les Commissions de terminologie pouvant être à l'origine d'événements linguistiques.

²¹ Ces marques ont été notamment étudiées dans *Langue française*, n° 73, *La reformulation du sens dans le discours*, Larousse, 1987.

²² Authier-Revuz, J. *et alii* (2003).

L'internet qui se définit comme un multimédia laisse une place prépondérante au mot qui est la clé de la recherche de l'information par les moteurs de recherche ou les annuaires. Le choix des mots-clés permettant d'obtenir des résultats pertinents explore (pour le concepteur des moteurs comme pour l'internaute) les réseaux sémantiques du lexique. Du côté de l'écriture, nous avons mis en évidence quelques modifications entraînées par l'utilisation du lien hypertextuel – notamment pour le discours de vulgarisation scientifique. Ce lien hypertextuel qui permet un retour sur les mots peut élargir l'ensemble des marqueurs de la modalisation autonymique.

ANNEXE

COMMISSION GÉNÉRALE DE TERMINOLOGIE ET DE NÉOLOGIE

Liste des termes, expressions et définitions adoptés de décembre 1997 à mars 1999

VOCABULAIRE DE L'INFORMATIQUE ET DE L'INTERNET

Liste du 16-3-1999. JO du 16-3-1999

1 - TERMES ET DÉFINITIONS [Extraits]

(Suivent en 2. des tables d'équivalence, anglais/français et français/anglais)

appliquette n. f.

Domaine : Informatique/Internet.

Définition : Petite application indépendante du matériel et du logiciel utilisés, qui est téléchargée depuis un serveur de la toile mondiale et qui est exécutée localement au sein d'un logiciel de navigation.

Note : Les appliquettes sont surtout employées dans le langage de programmation Java.

Voir aussi : logiciel de navigation, toile d'araignée mondiale.

Équivalent étranger : applet.

causette n. f

Domaine : Informatique/Internet

Définition : Communication informelle entre plusieurs personnes sur l'internet, par échange de messages affichés sur leurs écrans.

Voir aussi : internet.

Équivalent étranger : chat.

forum n. m.

Domaine : Informatique/Internet.

Définition : Service permettant des échanges et des discussions sur un thème donné : chaque utilisateur peut lire à tout moment les interventions de tous les autres et apporter sa propre contribution sous forme d'articles.

Voir aussi : article de forum.

Équivalent étranger : newsgroup.

fouineur n. m.

Domaine : Informatique/Internet.

Définition : Personne passionnée d'informatique qui, par jeu, curiosité, défi personnel ou par souci de notoriété, sonde, au hasard plutôt qu'à l'aide de manuels techniques, les possibilités matérielles et logicielles des systèmes informatiques afin de pouvoir éventuellement s'y immiscer.

(Source : Office de la langue française du Québec)

Voir aussi : pirate.

Équivalent étranger : hacker.

frimousse n. f. fam.

Domaine : Informatique/Internet.

Définition : Dans un message, association facétieuse de quelques caractères typographiques qui évoquent un visage expressif.

Note : 1. Les deux frimousses les plus connues sont : -) pour la bonne humeur et : - (pour le dépit, où les deux points représentent les yeux, le trait représente le nez et les parenthèses la bouche.

2. Le terme « binette » est recommandé au Québec. « Frimousse » doit être préféré à « binette ».

Équivalent étranger : emoticon, smiley.

hypertexte n. m.

Domaine : Informatique/Internet.

Définition : Système de renvois permettant de passer directement d'une partie d'un document à une autre, ou d'un document à d'autres documents choisis comme pertinents par l'auteur.

Équivalent étranger : hypertext.

internaute n.

Domaine : Télécommunications/Réseaux-Internet.

Définition : Utilisateur de l'internet.

Note : On rencontre aussi le terme « cybernaute ».

Équivalent étranger : cybernaut.

internet n. m. sg.

Domaine : Télécommunications/Réseaux-Internet.

Définition : Réseau mondial associant des ressources de télécommunication et des ordinateurs serveurs et clients, destiné à l'échange de messages électroniques, d'informations multimédias et de fichiers. Il fonctionne en utilisant un protocole commun qui permet l'acheminement de proche en proche de messages découpés en paquets indépendants.

Note : L'acheminement est fondé sur le protocole IP (Internet Protocol), spécifié par l'Internet Society (ISOC). L'accès au réseau est ouvert à tout utilisateur ayant obtenu une adresse auprès d'un organisme accrédité. La gestion est décentralisée en réseaux inter-connectés.

Équivalent étranger : Internet network, Internet, Net.

RÉFÉRENCES BIBLIOGRAPHIQUES

(Les adresses des sites ont été vérifiées en novembre 2005)

AUTHIER-REVUZ, J., DOURY, M., REBOUL-TOURÉ, S., Ed (2003), *Parler des mots – Le fait autonymique en discours*, Presses Sorbonne nouvelle, Paris.

BERTEN, F. (2000), *La notion de mot-clé et sa difficile application pédagogique*, Commission « français et informatique », <http://users.skynet.be/ameurant/francinfo/motcle/motcle.html>

BOSREDON, B., TAMBA, I., PETIT, G. Ed (2001), *Cahiers de praxématique – Linguistique de la dénomination* N°36, Publications Montpellier 3.

LE DIBERDER, A. (2001), « Sabir cyber », *Le Monde* du mardi, « supplément interactif », http://www.iclve.com/frenchclve/sabir_cyber.htm

Dictionnaire historique de la langue française (2000), Rey, A. et alii, dictionnaires Le Robert.

GUILHAUMOU, J. (2002), « Evènement discursif », « Evènement linguistique » in Charaudeau, P. et Maingueneau, D. Ed., *Dictionnaire d'analyse du discours*, Seuil.

HUMBLEY, J. (2001), « Quelques enjeux de la dénomination en terminologie », *Cahiers de praxématique – Linguistique de la dénomination* N°36, Publications Montpellier 3, 117-139.

KERBRAT-ORECCHIONI, C. (2003), *La norme en linguistique*, articles du séminaire « Normes, règles, régularités » des Ecoles doctorales SHS-Sciences humaines et sociales, et ECLIPS-Education, Cognition, Langages, Interactions, Psychologie, http://www.univ-lyon2.fr/article.php?id_article=653

MORTUREUX, M.-F. (1982), « Paraphrase et métalangage dans le dialogue de vulgarisation », *Langue française* N° 53, Larousse, 48-61.

MORTUREUX, M.-F. (1993), « Paradigmes désignationnels », *Semen* 8, Université de Besançon, 123-141.

PETIT, G. (2001), « Pour une conception lexicologique de la dénomination », *Cahiers de praxématique – Linguistique de la dénomination* N° 36, Publications Montpellier 3, 93-115.

REBOUL-TOURÉ, S. (2004), « Écrire la vulgarisation scientifique aujourd'hui », colloque *Sciences, Médias et Société*, 15-17 juin 2004, Lyon, ENS-LSH, http://sciences-medias.ens-lsh.fr/article.php?id_article=65

Recherche intelligente sur l'internet, Revue *Pour la science* N° 262, août 1999, <http://www.pourlascience.com>

Vocabulaire de l'informatique et de l'internet, J.O. du 16-3-1999, <http://www.education.gouv.fr/bo/1999/14/encart.htm>

Vocabulaire de l'internet, J.O. du 20 mai 2005, <http://www.culture.gouv.fr/culture/dglf/cogeter/20-05-05-internet.htm>

Vocabulaire du courrier électronique, J.O. n° 141 du 20 juin 2003, <http://www.culture.gouv.fr/culture/dglf/cogeter/20-06-03-courriel.htm>

Wikipédia, l'encyclopédie libre, <http://fr.wikipedia.org/wiki/Accueil>