



HAL
open science

Les sens de numérique – émergence d’emplois et dynamique du changement sémantique

Sascha Diwersy, Francesca Frontini, Agata Jackiewicz, Giancarlo Luxardo, Agnès Steuckardt

► **To cite this version:**

Sascha Diwersy, Francesca Frontini, Agata Jackiewicz, Giancarlo Luxardo, Agnès Steuckardt. Les sens de numérique – émergence d’emplois et dynamique du changement sémantique. 9èmes Journées Internationales de la Linguistique de corpus, Jul 2017, Grenoble, France. <hal-01568438>

HAL Id: hal-01568438

<https://hal.science/hal-01568438v1>

Submitted on 25 Jul 2017

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.



HAL Authorization

Les sens de *numérique* – émergence d’emplois et dynamique du changement sémantique

Sascha Diwersy , Francesca Frontini , Agata Jackiewicz , Giancarlo Luxardo et Agnès Steuckardt
Praxiling UMR 5267, CNRS - Université Paul-Valéry Montpellier 3
prenom.nom@univ-montp3.fr

1. Introduction

L’avènement des nouvelles technologies de l’information et de la communication s’accompagne de modifications substantielles de nos modes de vie et de nos manières de penser et de parler. Dorénavant, nous vivons à l’ère numérique, dans une société numérique. Nous manipulons des outils numériques avec plus ou moins de « bons sens numérique », bénéficions des services numériques dans des administrations, feuilletons des livres numériques, pratiquons une pédagogie numérique, participons à une démocratie numérique, via des Agoras numériques... La question du numérique se trouve ainsi au cœur d’enjeux sociétaux présentant des aspects interdépendants et parfois contradictoires. Les usages se multiplient, alors que les emplois établis (en rapport avec les nombres tout court) persistent.

Notre contribution s’intéresse à l’évolution diachronique des emplois de l’adjectif *numérique* en français contemporain. Elle est structurée de la façon suivante : la première partie est consacrée à la présentation du corpus utilisé dans notre étude et d’une méthodologie originale dédiée à la classification de corpus diachroniques. Dans la deuxième partie, nous exposons les résultats obtenus grâce aux traitements lexicométriques effectués pour proposer ensuite, dans la dernière partie, un bilan interprétatif et quelques perspectives de recherche.

2. Corpus et méthodologie d’exploitation lexicométrique

2.1. Corpus

Le corpus sur lequel est basée notre étude est composé d’articles de presse parus dans le quotidien *Le Monde* entre 1945 et 2015. Pour constituer cet échantillon, nous avons recueilli entre 70 et 80 numéros par année (soit de 1 à 2 numéros par semaine), pour un total de 671.943 textes et de 348.497.000 mots-occurrences. Le corpus a fait l’objet d’un étiquetage en parties du discours, d’une lemmatisation et d’une annotation en relations de dépendance syntaxique au moyen de la chaîne de traitement Bonsai (Candito *et al.* 2010) dans sa version qui fait appel à l’analyseur MaltParser (Nivre *et al.* 2006).

2.2. Méthodologie d’exploitation lexicométrique

Notre étude de l’évolution du sémantisme propre au lexème *numérique* est basée sur l’observation de la variation diachronique de son profil combinatoire, tel qu’il se manifeste à travers les collocations adjectivo-nominales impliquant *numérique* en tant qu’épithète. Nous avons donc partitionné le corpus en tranches de cinq ans (1945-1949, 1950-1954, 1955-1959, etc.), pour lesquelles nous avons respectivement établi un inventaire cooccurentiel des noms régissant l’adjectif *numérique*. Ces inventaires ont été répertoriés sous forme d’une liste, que nous avons soumise à un calcul de scores d’association, effectué dans la lignée de la méthodologie relevant de l’analyse dite collostructionnelle (Stefanowitsch & Gries 2003), au moyen du test exact de Fisher-Yates. Les résultats obtenus ont été enregistrés sous forme d’un lexicogramme (Heiden & Tournier 1998), dont le tableau suivant donne quelques extraits à titre illustratif :

Pour étayer notre analyse de l’évolution diachronique des emplois de l’adjectif *numérique*, nous avons appliqué à ce lexicogramme un procédé de périodisation automatique, à savoir la Classification Ascendante Hiérarchique par Contiguïtés (CAHC), dérivé de la méthode de classification proposée par Gries & Hilpert (2008 ; 2012) : Variability-based Neighbor Clustering (VNC). Cette méthode a été conçue pour le traitement de données lexico-statistiques relevant d’une variable d’ordre chronologique. La classification que nous avons mise en œuvre vise l’évolution des (scores de) dissimilarités entre les inventaires cooccurentiels adjacents au sein de la série constituée par les tranches temporelles issues

Mot-pôle	Collocatif	Sous-échantillon / Partie (tranche de 5 ans)	Co-fréquence	Score Fisher-Yates
numérique	importance	1945-1949	17	51,4177
...
numérique	supériorité	1950-1954	19	72,1411
...
numérique	commande	1975-1979	17	51,6906
...
numérique	télévision	1990-1994	38	84,7942
...
numérique	révolution	2010-2015	95	191,6682
...

TABLE 1 – Extrait du lexicogramme répertoriant les collocations adjectivo-nominales impliquant *numérique* en tant qu'épithète

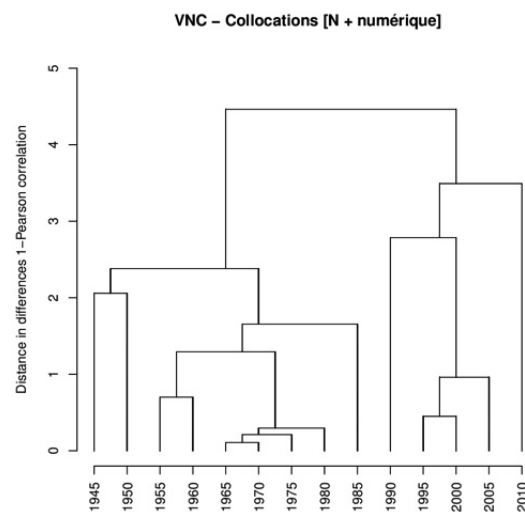


FIGURE 1 – CAHC appliquée aux collocations adjectivo-nominales impliquant *numérique* en tant qu'épithète

du partitionnement du corpus et par conséquent à mettre en évidence les discontinuités majeures que nous avons effectué en amont de l'analyse cooccurrentielle.

3. L'évolution diachronique des emplois de *numérique*

La CAHC appliquée au lexicogramme de *numérique* permet de produire un schéma combinant l'évolution dans le temps avec une classification hiérarchique (voir ci-dessous). Celle-ci met en évidence deux « coupes » majeures et trois périodes qui se dégagent successivement : 1945-1989, 1990-2009, 2010-2015 (le premier basculement illustré par le schéma intervenant sur 1990, le deuxième sur 2010).

Ces deux coupes peuvent être expliquées par l'extraction des collocatifs nominaux de la première tranche de la période, marquant la rupture avec la période précédente. En se limitant au premier décile (premiers 10

- pour la période 1990-1994 (marquant la rupture avec les périodes précédentes) :

télévision, compression, image, commande, TVHD, technologie, norme, définition, transmission, donnée, supériorité, infériorité

- pour la période 2010-2015 :

révolution, livre, économie, tablette, télévision, ère, technologie, plate-forme, version, radio, support, outil, art, fichier, photo, transformation, oubli, simulation, bouquet, écosystème, baladeur, caméra, image, infériorité, donnée, modèle, usages, contenu, univers, plateforme, bibliothèque, développement, mutation, édition, abonné, format, Interview, identité, média, Métropole, lecture, stratégie, offre, création, service, effet, trace, Syntec, information, appareil, innovation, culture

Si on part des quatre emplois majeurs¹ indiqués par *Le Petit Robert* 2016, on voit que les deux premières coupes signalées par la CAHC sont induites avant tout par des collocations impliquant l'emploi *numérique*₄ (cf. *Le Petit Robert* 2016, s.v. NUMÉRIQUE : « Se dit de la représentation de données, de grandeurs physiques au moyen de caractères tels que des chiffres (opposé à *analogique*), ainsi que des procédés utilisant ce mode de représentation (...) »), ce qui permet de dresser un bilan intermédiaire, qui consiste à constater d'une part, par rapport à la période 1990-1994, un « enracinement » évident de l'emploi *numérique*₄ renvoyant, par opposition à *analogique*, à une technologie et aux appareils qui utilisent celle-ci et d'autre part, pour la période à partir de 2010, une forte présence de collocations marquant différentes extensions métonymiques de cet emploi, dont la plus emblématique concerne la qualification des domaines et phénomènes sociétaux (*économie, art, univers, écosystème – révolution, ère, développement, mutation*) relatifs à cette technologie, à son utilisation et à ses produits. Le changement de collocation repéré dans l'usage de l'adjectif *numérique* à partir de 2010 correspond sans doute à une progression des usages sociaux d'internet. Il semble que cet emploi, qui apparaît, du point de vue du parcours sémantique diachronique et logique comme un emploi plutôt restreint, tende à devenir le sens prototypique de *numérique*. L'usage actuel de cet adjectif se caractérise à la fois par ce caractère massif et par ce brouillage sémantique.

4. Perspectives de recherche

L'évolution des emplois de l'adjectif *numérique* telle qu'on peut l'observer grâce à l'analyse collocationnelle que nous venons de présenter amène plusieurs questionnements.

D'un point de vue sémantique, en étudiant les classes de collocatifs nominaux, il y a lieu de se demander si l'évolution récente des emplois qui se manifeste à travers ces classes ne s'ouvre pas à des emplois subjectifs et en particulier axiologiques, comme le montrent des collocatifs tels que *fracture, vulnérabilité* ou encore *innovation*. A cet égard, il peut s'avérer très instructif d'étudier les emplois en extension de numérique en observant son fonctionnement dans le cadre de constructions relevant d'une grammaire locale telles que « N numérique de GN » (traces numériques des usagers, vulnérabilité numérique des voitures connectées, souveraineté numérique des pays, sabotage numérique des installations industrielles...). D'un point de vue sémantique, en étudiant les classes de collocatifs nominaux, il y a lieu de se demander si l'évolution récente des emplois qui se manifeste à travers ces classes ne s'ouvre pas à des emplois subjectifs et en particulier axiologiques, comme le montrent des collocatifs tels que *fracture, vulnérabilité* ou encore *innovation*. A cet égard, il peut s'avérer très instructif d'étudier les emplois en extension de numérique en observant son fonctionnement dans le cadre de constructions relevant d'une grammaire locale telles que « N *numérique* de GN » (*traces numériques des usagers, vulnérabilité numérique des voitures connectées, souveraineté numérique des pays, sabotage numérique des installations industrielles...*).

D'un point de vue syntaxique et morphologique, il semble pertinent de se poser la question de savoir si la dynamique du changement sémantique qu'on peut constater par rapport à l'extension des emplois de *numérique* va de concert avec des changements aux niveaux syntaxique (cf. la conversion vers l'emploi nominal de *numérique* adjectif) ou dérivationnel.

Au-delà de ces questions, d'autres explorations pourraient compléter le programme de travail que nous venons d'exposer, et notamment la prise en compte de productions langagières plus actuelles et de différents genres et registres (presse *vs* messages de forum *vs* textes scientifiques, etc.). De même, il sera indispensable d'entreprendre une étude contrastant *numérique* avec *digital, informatique* ou *télématique*.

1. Les emplois majeurs de l'adjectif *numérique* selon *Le Petit Robert* 2016 sont : *numérique*₁ : « Qui est représenté par un nombre, se fait avec des nombres » ; *numérique*₂ : « Qui concerne les nombres arithmétiques » ; *numérique*₃ : « Évalué en nombre » ; *numérique*₄ : « Se dit de la représentation de données, de grandeurs physiques au moyen de caractères tels que des chiffres (opposé à *analogique*), ainsi que des procédés utilisant ce mode de représentation (...) ».

Références bibliographiques

- Candito, M.-H., Nivre, J., Denis, P. & Henestroza Anguiano, E. (2010). Benchmarking of Statistical Dependency Parsers for French. *Proceedings of The 23rd International Conference on Computational Linguistics (COLING 2010)*. Beijing, China.
- Gries, S.T., Hilpert, M. (2008). The identification of stages in diachronic data : variability-based neighbour clustering. *Corpora* 3 (1), pp. 59-81.
- Gries, S.T., Hilpert, M. (2012). Variability-based neighbor clustering : a bottom-up approach to periodization in historical linguistics, in : Nevalainen, T., Traugott, E. (eds.), *The Oxford handbook of the history of English*, Oxford : Oxford University Press, pp. 134-144.
- Heiden, S., Tournier, M. (1998). Lexicométrie textuelle, sens et stratégie discursive, in : *Actes I Simposio Internacional de Análisis del Discurso*, Madrid.
- Le Petit Robert de la langue française* (2016). Paris : Dictionnaires Le Robert. [version en ligne]
- Nivre, J., Hall, J., Nilsson, J. (2006). MaltParser : A Data-Driven Parser-Generator for Dependency Parsing. *Proceedings of the fifth international conference on Language Resources and Evaluation (LREC 2006)*. Genoa, Italy.
- Stefanowitsch, A., Gries, S.T. (2003). Collostructions : Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics*, 8.2, 209-43.