

La mise en mots de la culture informationnelle et sa traduction dans les textes français relatifs à l'Education aux Médias et à l'Information

The putting into words of the information literacy and its translation into the French texts relating to Media and Information Education

Eric Delamotte¹, Odile Chenevez², Philippe Jeanne³

¹ Université de Rouen Normandie, UMR CNRS 6590 ESO, eric.delamotte@univ-rouen.fr

² ANR Translit, odile.chenevez@orange.fr

³ Université de Rouen Normandie, UMR CNRS 6266 IDEES, philippe.jeanne@univ-rouen.fr

RÉSUMÉ. La fondation relativement récente de l'enseignement de la Culture Informationnelle au sein des systèmes scolaires, contribue au processus de légitimation de celle-ci avec, notamment, la mise en place de l'EMI dans le système éducatif français. Nous avons émis l'hypothèse que les *éducations* aux trois cultures dites informationnelles (information literacy, digital literacy and media literacy) sont constituées de savoirs fragmentés que la culture numérique (dans ses pratiques socio-économiques et scolaires) pourrait bien intégrer en un récit commun. Nous avons constitué deux corpus (2013, 2016), afin de mettre en lumière les convergences et divergences discursives entre les trois cultures et leur prise en compte en France dans les textes institutionnels récents. L'intégration ou la convergence de ces cultures, au sein d'une doxa scolaire en construction, fait-elle apparaître de nouveaux équilibres ou déséquilibres ?

ABSTRACT. The relatively recent foundation of Information Literacy formation within school systems contributes to the legitimization process of it with, in particular, the introduction of the EMI (Media and Information Education) into the French education system. We made the assumption that the education of the three so-called information literacies (information literacy, computer literacy and media literacy) is made up of fragmented knowledge that the digital culture (in its socio-economic and educational practices) could well integrate into a common narrative. We have created two corpuses (2013, 2016) in order to highlight the discursive convergences and discrepancies between the three cultures and their inclusion in recent institutional texts in France. Does the integration or convergence of these cultures, within a school doxa under construction, reveal new balances or imbalances?

MOTS-CLÉS. Cultures informationnelles, Translittératie, Education aux Médias et à l'Information, Numérique Discours, Institutionnalisation, Doxa, Alceste.

KEYWORDS. Information Literacies, Transliteracy, Media and Information Education, Digital, Speech, Institutionalization, Doxa, Alceste.

Introduction

Léa (11 ans, France) : « *Il me faut une page Facebook, j'ai des copines qui en parlent, elles ont toutes des pages Facebook, elles, et elles disent que c'est bien, donc je dois essayer, et je me dis que ce serait bien pour avoir comme elles aussi, sinon je suis pas dans le coup quand elles en parlent* » (Cordier, 2015 : 122)¹.

Alicia (17 ans, USA) : « *Je pense tout simplement que (la technologie) est en train de redéfinir ce qu'il est acceptable que les gens montrent d'eux-mêmes. J'ai grandi avec la technologie, alors je ne sais pas ce qu'il était avant le boom des réseaux sociaux. Mais je pense qu'au lieu de passer notre temps à parler avec d'autres individuellement et partager avec eux des choses qui nous apparaissent comme privées, nous passons tout notre temps à mettre tout cela dans un seul module de*

¹ Cordier A. (2015), *Grandir connectés. Les adolescents et la recherche d'information*, Caen : C&F éditions.

communication où les gens peuvent venir et y avoir accès s'ils en ont envie. C'est plus pratique ». (boyd, 2016 : 142)².

Depuis plusieurs années le monde de la recherche se confronte au réel de notre société dorénavant en partie numérique. Certains chercheurs comme Anne Cordier ou danah boyd sont partis à la rencontre des jeunes pour les questionner sur leurs pratiques d'information et de communication. Ce type d'enquête permet de mieux comprendre ce que ressentent les adolescents devant les injonctions qu'ils subissent et les arts de faire qu'ils déploient quotidiennement. Loin des mythes ou des « unes » médiatiques, ces recherches commencent par écouter les adolescents, et restituer une Culture informationnelle entraînant de se constituer.

Si l'on peut décrire la Culture informationnelle, ses médiations avec l'entrelacement des logiques identitaires, éducatives, sociales et économiques, ses modes de légitimation à partir des pratiques, on peut aussi observer comment les institutions à leur manière dessinent une Culture de l'information « officielle ». Les paroles mêmes des jeunes invitent à étudier le socle normatif de la Culture informationnelle en gestation.

Dans cet article une première partie de cette présentation sera consacrée à une définition de la Culture de l'information et à sa traduction dans le domaine éducatif. Nous voudrions ainsi pointer quelques facteurs qui permettent d'appréhender l'enjeu de la « normalisation » de la Culture informationnelle dans la sphère éducative.

Dans une seconde partie, plus longue, nous préciserons notre démarche théorique et notre cadre méthodologie, car les questions ouvertes par cette perspective sont aussi nombreuses qu'importantes pour saisir les formations discursives, c'est-à-dire des configurations dynamiques, hétérogènes, composites qui constituent des unités de savoir (Le Marec, 2002). Elles concernent la fabrication, la circulation et le succès (ou insuccès) de notions, de valeurs et de pratiques sociales.

La troisième partie sera consacrée aux résultats de deux études réalisées avec le logiciel Alceste³.

1. La Culture de l'information : questionnements

Trop souvent encore, l'existence de la société de l'information est posée comme allant de soi. Pourtant le concept de société de l'information a fait l'objet de nombreux débats, parmi lesquels ceux relatifs à la dénonciation d'une « idéologie technique » (Wolton, 2009). Nous rappellerons à ce propos les propos fondateurs de Claude Bartz sur un impératif conceptuel : « *La société d'information, à mon sens, ne commencera à pouvoir être dite telle que lorsqu'il y aura – on aurait dit il y a quelques années un « supplément d'âme » - lorsqu'il y aura ce qu'on va appeler rapidement une « culture » qui lui correspond. Donc l'hypothèse de base que je vous propose est la suivante : pas de société d'information sans culture informationnelle* » (Bartz, 1998 : 76).

A l'instar de la « Culture de masse » des années 1960 décrite notamment par Edgard Morin, la Culture informationnelle n'est pas une forme de culture inférieure ou dégradée sur une échelle qui serait dominée par la technique et les médias. Elle n'est pas non plus, au sens ethnologique, une culture spécifique à un groupe particulier (elle n'est pas la forme contemporaine d'une « culture populaire » ou d'une culture « jeune ») ; mais elle est une Culture au sens anthropologique, c'est-à-dire un ensemble singulier de représentations du monde issu d'un mode spécifique de production et prétendant articuler

² boyd d. (2016), *C'est compliqué. La vie numérique des adolescents*. (traduction française de *It's complicated : the social lives of networked teens*, Yale University Press, 2014) Caen : C&F éditions.

³ Les deux études ont été réalisées dans le cadre de l'ANR TRANSLIT. Les chercheurs de ces deux études en dehors des auteurs de cet article sont : Alexandre Serres, Marlène Loicq, Cédric Fluckiger, Marie-Laure Compant Lafontaine, Ivana Ballarini.

les dimensions individuelles et collectives, réelles et imaginaires, de l'existence. Selon cette perspective, Milad Doueïhi, dans son ouvrage *Pour un humanisme numérique*, considère que nous sommes plutôt dans une situation de fait : la Culture de l'information est « le résultat d'une convergence entre notre héritage culturel complexe et une technique devenue un lieu de sociabilité sans précédent » (2011).

La question de la définition de la culture informationnelle mérite d'être abordée sous l'angle de l'éducation. En effet, la fondation relativement récente de l'enseignement de la Culture informationnelle au sein des systèmes scolaires, contribue au processus de légitimation de celle-ci tel qu'on peut le déduire et le reconstituer à partir des logiques relatives aux « normes » sociales et scolaires en émergence.

1.1. *L'analyse de la légitimation de la Culture informationnelle*

Aujourd'hui, la « capacité à maîtriser l'information » est devenue une préoccupation éducative à l'échelle internationale. On peut en percevoir quelques éléments à partir des déclarations de l'UNESCO qui tous s'appuient sur l'article 19 de la Déclaration universelle des droits de l'Homme de 1948⁴ pour promouvoir une éducation à l'information-communication : la Déclaration de Grünwald sur l'éducation aux médias (1982), la Déclaration de Prague relative à la compétence dans l'usage de l'information (2003), la Proclamation d'Alexandrie sur la maîtrise de l'information et l'apprentissage tout au long de la vie (2005), les 12 recommandations de l'Agenda de Paris (2007) ainsi que le Programme d'Education aux Médias et à l'Information pour les enseignants (2012) et, enfin, la « Déclaration de Qingdao », signée lors de la conférence de l'Unesco en juin 2015, est la première déclaration mondiale sur l'utilisation des TIC en éducation. Elle encourage leur utilisation et vise à lutter contre l'alphabétisme numérique. Il devient dorénavant impératif de former des personnes qui savent utiliser le numérique pour chercher, comprendre, synthétiser, communiquer et innover.

C'est un enjeu partagé dans les pays développés : tout le monde est d'accord pour convenir que l'éducation à l'information aux médias et au numérique est un problème crucial. Et la capacité des usagers à filtrer, évaluer la crédibilité des sources, etc. est au cœur de la maîtrise de l'information (Serres, 2012). Mais ce consensus et sa concrétisation s'opère à des rythmes et selon des modalités différentes dans la plupart des pays et concerne aussi bien tous les niveaux éducatifs.

En conséquence, il est particulièrement intéressant d'analyser ce processus d'agencement où la culture scolaire, le monde économique tout comme les pratiques sociales issues de la convergence des industries de la connaissance, de la communication et des loisirs, exercent leurs empreintes pour élaborer une « grammaire » info-communicationnelle (Delamotte, 2013).

Reste que l'hybridité ou la multimodalité des pratiques est un enjeu pour la recherche dans la mesure où elles invitent également à poser la question du territoire de la culture informationnelle (*Information literacy*), notamment au regard de cultures proches liées au domaine de l'informatique (*Digital Literacy*) ou à celui des médias (*Media Literacy*). Car pour certains, la maîtrise de l'information est considérée comme le champ d'étude principal qui comprend l'éducation aux médias, tandis que pour d'autres, la maîtrise de l'information n'est qu'une partie de l'éducation aux médias, qui est considérée comme le domaine d'étude principal⁵. Pour d'autres encore, il s'agit de trois champs issus de cultures différentes, à l'intersection non vide et qui se complètent.

⁴ « Tout individu a droit à la liberté d'opinion et d'expression, ce qui implique le droit de ne pas être inquiété pour ses opinions et celui de chercher, de recevoir et de répandre, sans considérations de frontières, les informations et les idées par quelque moyen d'expression que ce soit. »

⁵ En témoigne : Monique Lebrun, Nathalie Lacelle et Jean-François Boutin (dir), *Littératie médiatique multimodale. De nouvelles approches en lecture-écriture à l'école et hors de l'école*, Québec, Presses de l'Université du Québec, 2012.

1.2. La culture informationnelle et sa traduction en EMI dans le contexte éducatif international

L'essor de l'*Education aux Médias et à l'Information* (EMI) et son évolution à l'échelle mondiale sont étroitement liés à plusieurs facteurs très différents, que l'on retrouve cependant dans la plupart des régions du globe :

- l'explosion des « nouvelles technologies » et notamment d'Internet qui bouleverse les modes de recherche et d'exploitation de l'information. Ce facteur en constante évolution, influe largement sur le travail des élèves et des étudiants et des formateurs.

- la prise de conscience de développer une capacité à agir sur les réseaux numériques d'une manière responsable, prudente et éthique, en accord avec ses valeurs et ses identités

- une réforme de l'enseignement intervenue dans de nombreux pays occidentaux (Europe, Australie notamment) pour tenter de s'adapter à de nouveaux publics et à des besoins économiques. Ces réformes insistent sur les compétences à développer chez les élèves (c'est notamment le cas en France avec le Socle commun de connaissances, de compétences et de culture).

- un accent plus ou moins fort est mis dans la plupart des pays sur la nécessité de se former tout au long de la vie et l'utilité de cette compétence dans les critères d'employabilité aux Etats-Unis et en Europe notamment.

Les rapports européens *Les compétences clés pour l'apprentissage tout au long de la vie* (2004) puis *Les compétences clés dans un monde en mutation* (2009) et enfin plus récemment, en France, le Rapport du Conseil National du Numérique intitulé *Citoyen d'une société numérique* (2013) font état de la nécessité de développer la compétence à chercher, à évaluer et à utiliser l'information. La « compétence numérique » est définie comme compétence « transversale », au même titre que l'apprendre à apprendre et l'esprit d'entreprise⁶. On constate, pour le moins, la construction de standards cherchant à évaluer et apprécier les compétences atteintes par les acteurs observés⁷.

En France, comme d'autres domaines que l'on classe dans les « éducations à », l'Education aux Médias et à l'Information (EMI), les dispositifs et les didactiques ou configurations didactiques qui en résultent sont l'occasion de questionnements nouveaux ou renouvelés. Si les « éducations à » sont clairement centrées sur des pratiques sociales à visée participative, en prise avec la complexité d'un monde incertain et mouvant, quelles implications didactiques en résulte-t-il ? Selon quelles incarnations ? Selon quelle spécificité pour chacune d'entre elles ? Et selon quelles modalités d'évaluation ?

Sous un seul terme générique (EMI), on rapproche deux domaines. D'une part, la maîtrise de l'information met l'accent sur l'importance de l'accès à l'information, son évaluation et son utilisation éthique. D'autre part, l'éducation aux médias met l'accent sur la capacité de comprendre les fonctions des médias, d'évaluer la manière dont ces fonctions sont exercées et de faire usage de ces médias de façon rationnelle pour s'exprimer. Cependant une nouvelle composante de la Culture de l'Information s'impose, car nos comportements, nos relations à l'espace, aux temps et aux autres sont impactés par les spécificités de la médiation instrumentale propres aux machines à communiquer. La *digital literacy*

⁶ Pedró Francesc, *The new millennium learners: challenging our views on ICT and learning*, OECD-CERI, 2006.
<http://www.oecd.org/dataoecd/1/1/38358359.pdf>

⁷ Parmi eux, il faut noter le « Kit de l'éducation aux médias », dans le cadre de l'UNESCO, le développement d'indicateurs d'éducation à l'information (avec IFLA) ou encore la publication d'une première cartographie des politiques d'éducation aux médias dans le monde (Frau-Meigs et Torrent, 2009) et, enfin, les travaux français de l'Equipe de Recherche Technologique en éducation (ERTé) « *Culture informationnelle et curriculum documentaire* » pilotée par Béguin Annette, ayant donné lieu à une publication sous la direction scientifique de Chapron Françoise et Delamotte Eric (2010).

tente de se différencier d'une *computer literacy* d'essence davantage informatique et basée sur de la programmation. Si l'ordinateur apparaît comme le principal support des activités informatiques scolaires, celui-ci sert le plus souvent à la science informatique entendue dans ses dimensions algorithmiques ou de programmation qu'à sa dimension technique. De plus, pour les chercheurs, la double dimension des technologies numériques reste une question sensible : tantôt on insiste sur leur existence comme objets (objets d'apprentissage en eux-mêmes), tantôt on souligne leur existence comme outils (moyens de rapport à des objets de savoir, de communication, etc.), jusqu'à ce que cette distinction soit recouverte par le concept de « ressource » (Fluckiger et Hétier, 2014)⁸.

2. Orientations théoriques et méthodologiques

Il n'est donc pas étonnant que ces interrogations majeures s'imposent à l'institution scolaire qui dit-on, doit s'adapter, au plus vite et au mieux, à l'expansion d'une culture numérique. Le danger serait de poser le numérique comme un commencement absolu, faisant largement table rase du passé, au lieu d'essayer de comprendre comment il modifie, par des ajouts mais aussi des pertes, ce qui a toujours existé dans la vie sociale : des outils, des langages, des symboles, des relations, des institutions, des créations.

Pour qualifier les techniques dites « numériques »⁹, la plupart des discours contemporains d'ordre médiatique, publicitaire ou politique empruntent fréquemment leur vocabulaire au registre du « changement »¹⁰. En conséquence notre premier geste en matière de recherche doit consister à déconstruire le « discours d'escorte »¹¹ porté par différents acteurs au nom de la société. La notion de « discours d'escorte » est ici employée de façon large, elle désigne tout type de discours qui « accompagne » les objets, dispositifs techniques ou pratiques. Aussi le « discours d'escorte » appartient aux « formations discursives » au sens où l'entendait Michel Foucault (*L'Archéologie du savoir*, Gallimard, 1969) où ces discours ont une force agissante et « performative » sur le social. D'une façon générale, pour Paul Ricoeur, tout récit scientifique, politique ou autre qui raconte « ce qui se passe » institue du réel. Pourtant, le réel représenté ne correspond pas seulement au réel. Il cache, derrière la figuration du présent, un passé et un futur qui l'ordonnent¹². Ainsi, pour Michel de Certeau, « des récits marchent devant les pratiques sociales pour leur ouvrir un champ »¹³.

Pour autant, plus encore que la définition à donner au contexte d'énonciation des discours ambiants, c'est la notion de « sens commun » elle-même qui fournira le cadre de notre interrogation, ainsi que les modalités spécifiques de sa construction dans le discours des institutions.

⁸ Cédric Fluckiger & Renaud Hétier (coord), « Des élèves et des savoirs à l'ère numérique : regards croisés » *Recherches en Education* n°18 Janvier 2014

⁹ Sur l'usage du terme « numérique » et les enjeux liés aux choix terminologiques, voir Emmanuel Souchier, « La *lettrure* à l'écran. Lire & écrire au regard des médias informatisés », *Communication & langages*, n° 174, décembre 2012, p. 85-108.

¹⁰ Ces discours de généralité sont pour la plupart des discours sur le changement, que celui-ci soit attendu ou redouté.

¹¹ Pour Gérard Genette, un texte se présente rarement à l'état nu, sans le renfort de l'accompagnement d'un certain nombre de productions tels que les titres, sous-titres, préfaces, notes, prières d'insérer, et bien d'autres entours moins visibles mais non moins efficaces, qui sont, pour le dire trop vite, le versant éditorial et pragmatique de l'oeuvre littéraire. Gérard Genette nomme ce « discours d'escorte qui accompagne tout texte » le paratexte. Par extension, un discours d'escorte accompagne une analyse littéraire, un choix d'extraits ou une sélection de textes. Son objectif est de rendre la matière accessible au lecteur, d'expliquer, de relier, de créer des liens, etc.

¹² Paul Ricoeur, *L'idéologie et l'utopie*, Paris, Le Seuil, 1997.

¹³ Michel de Certeau, *L'invention du quotidien I Arts de faire*, Paris, Gallimard-Folio, 1998, p. 185.

Dans le deuxième temps de notre recherche, nous posons comme hypothèse que les textes institutionnels relatifs à l'EMI dépassent en les englobant deux modèles relatifs à la fabrication des savoirs scolaires : celui qui se réfère à un modèle de simple transformation des savoirs des disciplines scientifiques en savoirs à enseigner¹⁴ ; celui qui renvoie aux pratiques sociales de références pour fonder un curriculum formel¹⁵. Les éléments de cet englobement doivent être toutefois discutés.

L'analyse des textes institutionnels, considérés comme de lieux de discours particulièrement composites qui concentrent du sens commun, des croyances, des savoirs scientifiques et des pratiques permet d'éclairer, nous l'avons dit, une réalité de la Culture informationnelle. Nous avons émis l'hypothèse que les *éducations aux trois cultures* dites informationnelles (au sens de *knowledge, news* et *data*) sont constituées de savoirs fragmentés que la culture numérique (dans ses pratiques socio-économiques et scolaires) pourrait bien intégrer en un récit commun. À ce propos, Antonio Gramsci écrivait : « La philosophie d'une époque n'est pas la philosophie de tel ou tel philosophe, de tel ou tel groupe d'intellectuels, de telle ou telle grande division des masses populaires : c'est la combinaison de tous ces éléments qui tend vers une direction donnée et dont l'aboutissement devient règle d'action collective, devient « histoire » concrète et complète (intégrale) ». La réflexion porte, dans ce second temps, sur les éléments qui constituent la mise en récit d'une norme scolaire en émergence.

2.1. Le modèle de la doxa

Le concept de « sens commun » indique la communauté d'un sens partagé, c'est-à-dire préalablement donné, sinon accepté, le plus souvent sous le mode de l'évidence. Ce concept s'avère fécond car il permet d'apparier, dans l'étude des objets soumis à cette perspective, le point de vue du savoir commun (aspect épistémique) et celui des croyances communes (aspect idéologique du sens commun). Par là, le discours institutionnel rejoint d'abord les questions posées par l'épistémologie sociale et relatives à la confiance épistémique, que cette confiance en l'autorité d'une source s'appuie sur des critères supposés rationnels ou qu'elle repose sur une « déférence » socialement construite (Origgi, 2004 ; 2008).

Afin, d'analyser la nature doxique du corpus, nous reprendrons à notre compte quatre propositions théoriques formulées par George Elia Sarfati (2002), à savoir :

- la fonction de codification se conçoit sous le rapport d'une pratique de légitimation et de légalisation. Au sein d'une fonction de *codification*, les sources se réfèrent avant tout à elles-mêmes.

- la fonction de *validation* prend en compte deux pratiques. La pratique qui consiste à caractériser le domaine de définition ainsi qu'à configurer les lignes de développement possibles de l'axiologie instituée. La pratique de limitation consiste quant à elle, à prendre en charge la mémoire discursive de l'institution.

¹⁴ Modèle répandu que l'on a trop souvent tendance à confondre avec la théorie de Yves Chevallard (1985/1991). *La transposition didactique. Du savoir savant au savoir enseigné*, Paris, La Pensée Sauvage. Pour cet auteur, le savoir à transposer pour l'enseigner dans une institution donnée n'est en effet pas restreint aux savoirs dits « scientifiques », mais inclut tout type de savoir concernant toute activité humaine. Il peut s'agir de planter un clou, de construire une argumentation ou de résoudre une question mathématique donnée... Les savoirs étant organisés en ensembles de « praxéologies » avec leurs 4 composantes : type de tâche, technique, technologie, théorie. Voir par exemple : Yves Chevallard (2010) *La didactique, dites-vous ?* Education & didactique, 4.1, 139-146. http://yves.chevallard.free.fr/spip/spip/article.php3?id_article=181

¹⁵ Jean Louis Martinand, J.-L. (1981). Pratiques sociales de référence et compétences techniques. À propos d'un projet d'initiation aux techniques de fabrication mécanique en classe de quatrième. In Giordan A. & Martinand J.-L. (dir.) *Diffusion et appropriation du savoir scientifique : enseignement et vulgarisation*. Actes des Troisièmes Journées Internationales sur l'Education Scientifique, Paris, Université Paris 7, (p. 149-154).

- La fonction de *différenciation* consiste principalement dans une opération discursive de séparation et de démarcation de l'axiologie instituée, relativement à d'autres axiologies connexes ou concurrentes. Elle est principalement une opération de tri entre les différents possibles et à spécifier des frontières.

- La fonction de *pérennisation* vise à assurer la viabilité ainsi que la recevabilité, à long terme, de la doxa instituée.

2.2. La scénographie institutionnelle

C'est dans le contexte étroit d'une scénographie institutionnelle particulière qu'il faut envisager la construction d'une doxa. Le discours des institutions semble le plus souvent sous-tendu par cet ethos des « locuteurs dissimulés », mis en évidence par Ruth Amossy (2010). En effet, l'effacement énonciatif sied aux responsables publics, car les institutions qui sont supposées parler d'une seule voix, s'accommodent mal du *dissensus* interne ou d'une polyphonie mal orchestrée.

Les institutions vont parfois jusqu'à récuser l'existence même du *dissensus* dans l'espace social, réduisant ou démentant les enjeux politiques qui sous-tendent la production et la circulation des discours au profit d'une lecture réputée rationnelle de la réalité. Roland Barthes, dans ses *Mythologies* [1957], avait su lui aussi disséquer le phénomène par lequel les fantasmes d'une époque s'offrent au public sous la forme fallacieuse d'essences universelles, et par là même naturalisent l'histoire, c'est-à-dire la niant.

Les institutions publiques tendent à un lissage de la conflictualité qui nous paraît caractéristique des discours institués. La force de l'évidence est alors le moyen d'étayer l'autorité et/ou d'imposer un cadre argumentatif naturalisé. Mais l'étude du lexique et des marques énonciatives montre sans peine la subjectivité et la polarisation axiologique qui traversent ces discours.

3. Analyses

Dans le cadre du projet de recherche ANR « Translit » avons avancé que les *éducations aux trois cultures* dites informationnelles (au sens de *knowledge*, *news* et *data*) sont constituées de savoirs fragmentés que la culture numérique (dans ses pratiques socio-économiques et scolaires) pourrait bien intégrer en une « grammaire » commune. Les phénomènes d'hybridation et d'articulation entre les 3 cultures concernées, dans leur transposition scolaire enseignable, pourraient en effet se laisser envahir par une « hégémonie impérialiste » de l'une ou l'autre d'entre elles (Serres 2012). Seule une analyse permet d'observer l'harmonie de l'intégration, et de souligner les éventuelles occultations ou absorptions invalidantes. Nous avons donc privilégié une approche d'analyse textuelle statistique, à l'aide du logiciel Alceste, sur des corpus de textes représentatifs des *discours d'influence* (Corpus 1) de chacune des 3 cultures d'une part, et de discours convergents ou supposés convergents, d'autre part.

Dans un second temps, nous avons porté nos efforts dans l'analyse de « normes » scolaires en construction relatives à l'éducation aux médias et à l'information. Il s'agissait principalement d'analyser les processus de validation « interne » puisque les productions discursives visent, en un premier temps, à décréter la pertinence d'un répertoire partagé préconisé et défendu. Précisément il s'agit de se pencher sur une mise en ordre scientifique et éducative qui permet de fabriquer une réduction simplifiée de l'*Education aux Médias et à l'Information* mais elle est inséparable d'un travail d'arbitrage où des chercheurs, des experts, des administrateurs, des groupes d'intérêts ont négocié et pesé sur la définition de l'EMI et sa mise en œuvre.

3.1. Le corpus 1

Un premier corpus de textes a été constitué en 2013 dans le but d'explorer les convergences et divergences entre les 3 cultures concernées. On parlera désormais de 3 *domaines*, désignés ainsi : éducation à l'info-doc, éducation à l'info-médias et éducation à l'info-data. Il s'agissait bien sûr de

prendre en compte les multiples embûches inhérentes à un tel exercice, en se dotant d'une méthodologie rigoureuse et de critères précis.

3.1.1. *Quelle méthodologie générale ?*

Le travail d'analyse textuelle avec Alceste a été envisagé par le groupe sur un premier corpus exploratoire, homogène et de petite taille, divisé en trois parties correspondant aux trois champs concernés. Dans un second temps, il était prévu une deuxième étape sur un corpus plus large et plus hétérogène. Cette tâche s'est transformée en l'étude du corpus de textes institutionnels.

Trois principes simples ont été dégagés pour guider la constitution du premier corpus :

- **la comparabilité des trois parties du corpus** : il s'agissait de constituer des mini-corpus à peu près équivalents, à la fois en taille, en types de textes et de sources ;

- **la cohérence interne** : l'ensemble du corpus devait viser une relative cohérence, en termes de types de documents, d'auteurs, de sources et d'approches ;

- **la définition de variables et de critères communs** : pour choisir les textes, un travail important de réflexion a été mené par le groupe sur la détermination des variables pertinentes et le choix de plusieurs critères, à la fois objectifs et précis.

Deux problèmes méthodologiques, classiques dans ce type d'analyses, ont été rapidement identifiés :

- **la représentativité** : les linguistes le disent eux-mêmes : "*les corpus textuels (qui) ne peuvent aspirer ni à l'exhaustivité ni même à la représentativité, (et qui) concentrent toujours des données attestées puisqu'on ne saurait fabriquer artificiellement un texte pour prétendre en appréhender le sens*". Ainsi, constituer un corpus représentatif des trois domaines était évidemment impossible, d'autant plus avec le choix d'un corpus restreint, de petite taille. Face à l'abondance des textes publiés au fil des années dans chacun des trois champs, prétendre vouloir identifier un petit ensemble de textes représentatifs aurait relevé d'une grave illusion scientifique et méthodologique. Aussi notre premier corpus ne prétend pas, évidemment, "représenter" à lui seul l'ensemble diversifié des textes issus des trois cultures de l'information. Il reflète, tout au plus, une certaine approche de ces trois cultures.

- **la subjectivité** : comment limiter au maximum la subjectivité des choix des textes ? Si l'affirmation d'une certaine vision des trois cultures de l'information (qui ne peut être développée ici) peut être assumée et revendiquée par le groupe, au moins dans la constitution globale du corpus, nous avons en revanche tenté de limiter, autant que faire se peut, notre subjectivité, nos choix personnels, dans la sélection précise de tel ou tel texte. (*extrait du rapport de recherche*)

Ce premier corpus, établi en 2013, contient 45 textes soigneusement choisis et équilibrés pour répondre à la fois à une exigence de représentativité des domaines comme de la qualification reconnue de « textes d'influence ». La valeur des résultats obtenus par l'analyse d'Alceste est bien évidemment indissociable d'un consensus quant à la validité du corpus sur ces critères. « L'effet corpus » prend ici une place reconnue et jamais occultée dans l'analyse des résultats.

Le choix de répartition fut de 15 textes pour chacun des domaines, avec, pour chacun, 5 textes en provenance de chacun des lieux d'énonciation (5 textes institutionnels, 5 textes scientifiques et 5 textes professionnels)¹⁶.

Ainsi en ce qui concerne les variables traitées par le logiciel, outre les auteurs et les dates de publication on retiendra ici particulièrement la variable « Lieu », qui définit le lieu d'énonciation, pouvant prendre 3 valeurs : institutionnel, scientifique ou professionnel.

¹⁶ Les références des textes du corpus 1 sont accessibles dans l'annexe 1 du rapport de recherche de la Tâche 1 de l'ANR Translit disponible sur collection HAL "Cultures de l'information".

3.2. Le corpus 2

En 2016 une nouvelle question de recherche a émergé, en lien direct avec la première : Quelle convergence des trois cultures de l'information existe vraiment dans les textes institutionnels depuis 2013 ? Autrement dit : Quels sont les éléments translittéraciques dans les textes officiels récents ?

Guidés par le modèle de G-E Sarfati (2002) nous avons notamment cherché à tester la part de canonicité de textes institutionnels récents, qui se veulent convergents et donc translittéraciques, et leur capacité à rendre compte des doxas issues des 3 cultures. Nous recherchions notamment des marqueurs de fonctions de validation et de pérennisation qui rendent possible la légitimation d'un discours. Dans cette perspective nous avons resserré la question de recherche : Comment les discours "spécifiques" des 3 cultures se retrouvent-ils dans les discours institutionnels ? Quels éléments de légitimation - issus de 3 cultures - apparaissent dans la nouvelle mise en récit ?

Si le critère principal du choix des textes restait le numérique comme principe englobant, avec la présence d'une forte dimension éducative, il s'agissait de se focaliser sur les seuls textes institutionnels, de prendre les programmes actuels et les textes prescriptifs et de recommandations et retrouver les influences scientifiques, la présence de termes issus des trois domaines, les prescriptions en matière de numérique...

35 textes ont ainsi été rassemblés selon les critères suivants : le corpus ne devait contenir que des textes officiels, provenant de différentes sources institutionnelles françaises (Ministère, Rectorats, organismes officiels...). Les textes devaient être à jour et toujours en vigueur, et la date de début du corpus était récente (2012). Ainsi la variable « Lieux (d'énonciation) » était réduite à un seul critère, celui des textes institutionnels.

La variable la plus importante, qui devait structurer ce corpus, était celle de la nature ou de la force de l'injonction. Si la quasi-totalité des textes choisis était de nature injonctive ou incitative (pour reprendre la typologie des textes selon leur fonction d'énonciation), de quel type était cette injonction ? Nous avons ainsi réparti les textes autour d'une échelle de quatre critères, selon la force plus ou moins grande de l'injonction : les textes de « Prescription » (ayant un caractère obligatoire, comme une loi, un décret, etc.), puis de « Préconisation » (comme une circulaire), les textes de « Recommandation » (comme certains rapports, les dossiers pédagogiques, etc.) et les textes de « Médiatisation » (comme les dossiers de presse), qui ne sont plus, à proprement parler, des textes « injonctifs »¹⁷.

Un premier corpus de textes a été constitué en 2013 dans le but d'explorer les convergences et divergences entre les 3 cultures concernées. On parlera désormais de 3 *domaines*, désignés ainsi : éducation à l'info-doc, éducation à l'info-médias et éducation à l'info-data. Il s'agissait bien sûr de prendre en compte les multiples embûches inhérentes à un tel exercice, en se dotant d'une méthodologie rigoureuse et de critères précis.

3.3. Le logiciel Alceste

L'objectif général de cette méthodologie vise, comme le précise son auteur, à « déterminer comment sont organisés les éléments qui constituent un texte » (Reinert, 1993). Il s'agit de « réduire l'arbitraire de la description du corpus en mettant en évidence ses régularités, ses symétries cachées » (Thom, 1974, 21). Alceste met l'accent sur les ressemblances et dissemblances du vocabulaire et rend compte de sa distribution dans les propositions qui constituent le texte étudié.

A travers des calculs statistiques le logiciel ALCESTE va mettre en évidence les particularités lexicales qui ponctuent le corpus. Après avoir procédé à une segmentation du texte, le logiciel repère les phénomènes de co-occurrence et procède à une mise en évidence de champs sémantiques relatifs à

¹⁷ Les références des textes du corpus 2 sont accessibles dans l'annexe 1 du rapport de recherche de la Tâche 1 de l'ANR Translit disponible sur collection HAL "Cultures de l'information"

ce qui est lexicalement commun à travers une classification descendante hiérarchique (désormais CDH). Les mécanismes statistiques s'appuient sur le repérage de « formes » graphiques, objets statistiques, indépendants du sens.

A partir d'une analyse factorielle des correspondances (désormais AFC), les unités sémantiques, issus de la CDH, sont positionnés sur un plan à deux dimensions permettant de visualiser d'une part ce qui est commun mais également de mettre en évidence des différences ainsi que les relations qu'entretiennent les classes entre elles. C'est la mise en évidence de « facteurs » qui structure le texte et associant ou dissociant ses parties. Ainsi, l'AFC renseigne sur les rapports de proximité ou d'éloignement des différentes parties du corpus caractérisées par la CDH, les unes par rapports aux autres (Marchand, 1998 ; 52).

La construction des champs lexicaux de contexte similaire, dénommées « classes », s'opère à partir des termes du lexique co-occurents, dont la significativité s'exprime par le calcul statistique du Khi^2 . Plus celui-ci sera élevé, plus la forme graphique repérée aura contribué à la construction de la classe. Ainsi constituées, elles rendent compte de l'organisation formelle du texte. Bien que reflet de l'organisation formalisée du corpus ces univers lexicaux offrent une palette d'interprétations sémantiques riche de ce que représente l'objet d'une part, mais également riche de l'interprétation qu'en fera l'analyste et qui sera fonction de sa propre identité et de son propre intérêt, d'autre part. De ce fait, le vocabulaire d'un énoncé devient une trace, une référence, une activité relevant d'une intention de sens du sujet-énonciateur. Ainsi, à travers l'interaction de l'individu et du discours, se dessinent des représentations sociales qui expriment le rapport que les uns et les autres entretiennent avec le monde.

A partir de ces calculs statistiques, le logiciel ALCESTE va décliner tout un panel d'interfaces graphiques qui doit permettre, au chercheur d'interpréter le plus justement possible, les résultats de l'analyse.

Dans la construction de notre interprétation, nous porterons un regard particulier à l'AFC. A l'intérieur d'une représentation graphique orthonormée, cette interface nous permet de visualiser l'articulation des différents mondes lexicaux, les uns par rapport aux autres. Ainsi, chaque terme est projeté sur les deux axes du plan en fonction d'une abscisse et d'une ordonnée, facteurs structurants, tandis que les champs lexicaux sont représentés sous forme de « patatoïdes ». La lecture est simple : plus les termes de lexique se regroupent auprès de l'intersection des axes, plus la nature du discours est neutre ; plus ils s'éloignent de l'intersection et plus la nature du discours se spécifie. Si les sphères lexicales s'enracinent toutes à l'intersection des axes (tout discours aussi spécifique soit-il, partage du vocabulaire avec les autres discours) il est intéressant de connaître ce qui les discrimine. Ceci est d'autant plus important que cette représentation permet en même temps d'étiqueter les quatre plans que dessine l'axe horizontal et vertical et doit permettre à l'analyste de nommer ces axes, véritables « fils conducteurs » structurants le discours.

L'analyse ALCESTE permet au chercheur d'ajuster sa focale sur des termes lexicaux spécifiquement présents ou absents. En effet, la première interface, « réseau de forme » permet de sélectionner un terme et de voir avec quels autres termes entretient-il, plus ou moins, une relation sémantique. Dans les dernières versions du logiciel, cette interface a été extrapolée en permettant de visualiser l'ensemble des mots d'une classe ou d'un individu (pris au sens statistique du terme) et de voir comment, plus ou moins, ils s'interconnectent. C'est alors un ensemble de relations sémantiques qui se dessinent et qui permet de distinguer des polarités lexicales caractérisées par des relations sémantiques extrêmement resserrées.

Cependant, aussi fort que soit notre désir d'objectivité, le seul fait de travailler sur le sens, impose l'engagement du sujet quelle que soit sa discipline. Aussi, si le logiciel, à travers une série de calculs statistiques, offre à l'utilisateur des résultats conséquents, ceux-ci ne feront sens qu'à travers

l'engagement du chercheur c'est-à-dire sur la manière de développer une interprétation empreinte d'une visée scientifique éthique.

Terminologie retenue pour la présentation de cette recherche :

Forme : chaîne de graphèmes analysés par Alceste.

Mot : unité graphique et phonologique pouvant être mono ou polysémique.

Terme : le terme ou unité terminologique est l'unité signifiante constituée d'un mot (terme simple) ou de plusieurs mots (terme complexe), qui désigne une notion de façon univoque à l'intérieur d'un domaine (Rey, 1992).

Il s'agit ici des termes spécifiques, propres à chaque domaine, et choisis par les experts avec un sens précis.

3.4. Résultats et interprétations du corpus 1

Parmi les analyses faites sur le corpus 1 la technique de la CDH a permis de distinguer 7 classes, organisées en arborescence. Les classes 1, 4 et 5 se distinguant en premier lieu des classes 3, 2, 6 et 7, comme les groupes de formes le plus différencié les uns des autres, puis les autres sous-ensembles de classes, puis les classes jusqu'à constituer ces listes de formes dont la cohérence est significative, tant par la présence de certaines formes que par l'absence d'autres (figure 1).



Figure 1. Le tableau des présences/absences de la CDH, définissant les classes du corpus

Alceste ne voit que des formes (suite de graphèmes séparés par des blancs) là où l'expert est capable de donner un sens générique à ces listes de mots. C'est ainsi qu'il peut poser une étiquette sur chacune des classes et groupes de classes de ce tableau.

Sur ce principe nous avons obtenu une arborescence des univers sémantiques du corpus, étiquetée par le groupe expert que constituait l'équipe de recherche (figure 2).

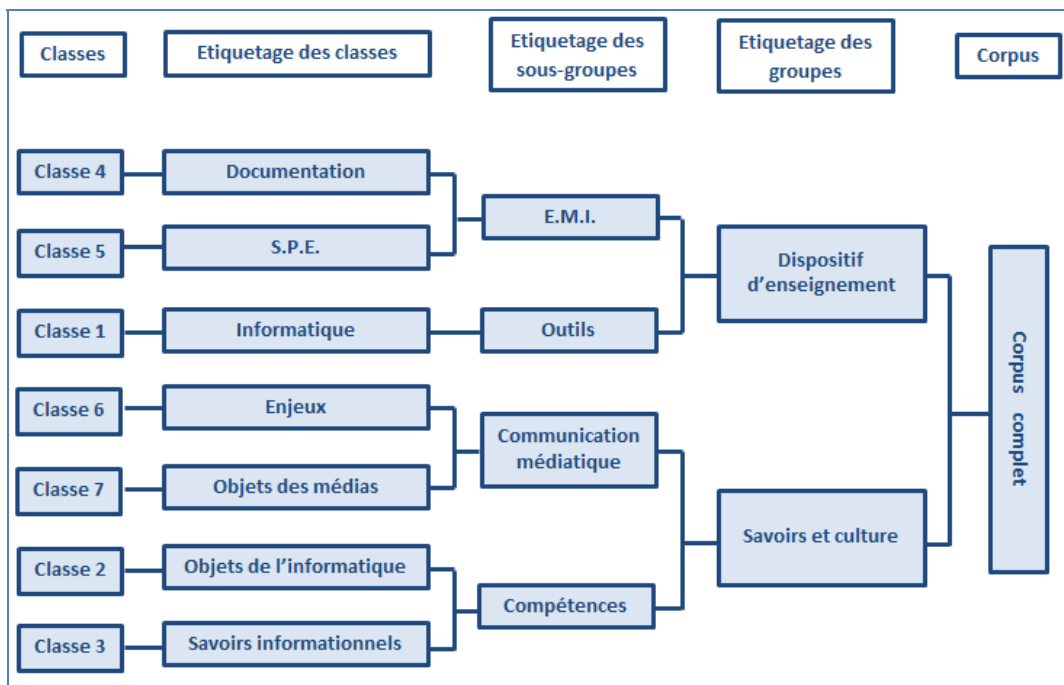


Figure 2. Organisation des univers sémantiques du corpus 1

On voit assez nettement dans cette figure que les thématiques distinguées dans le corpus correspondent presque hermétiquement aux domaines traités. Ainsi les classes 1 et 2 dessinent le domaine de l'info-data, la classe 4 et une grande partie de la classe 3 celui de l'info-doc et les classes 5, 6 et 7 celui de l'info-média. On voit aussi que chacun des domaines a un pied dans le groupe « dispositifs d'enseignement » et un pied dans le groupe « savoirs et culture ».

Toutes ces classes ne sont d'ailleurs pas d'égale importance en volume : la taille des cercles de couleur, au-dessus de chaque classe de la figure 1, nous renseigne sur ce fait. Les classes 3 et 4 étant les plus volumineuses en nombre d'unités textuelles représentées (respectivement 23 et 17 % des UCE). La plus petite étant la classe 7 concentrant seulement 9% des UCE du corpus.

De la même façon, nous avons pu réaliser une AFC « étiquetée »¹⁸ qui permet de faire apparaître une cartographie des classes (Fig. 3). Plus on est au centre de la carte plus les termes sont consensuels, convergents ; plus on s'éloigne du centre plus les formes ainsi désignées seront propres à leur classe et divergentes par rapport à l'ensemble du corpus.

¹⁸ Alceste place les nuages de points liés à la CDH, donc aux classes, le groupe expert définit les étiquettes d'après le sens des mots ainsi ordonnancés.

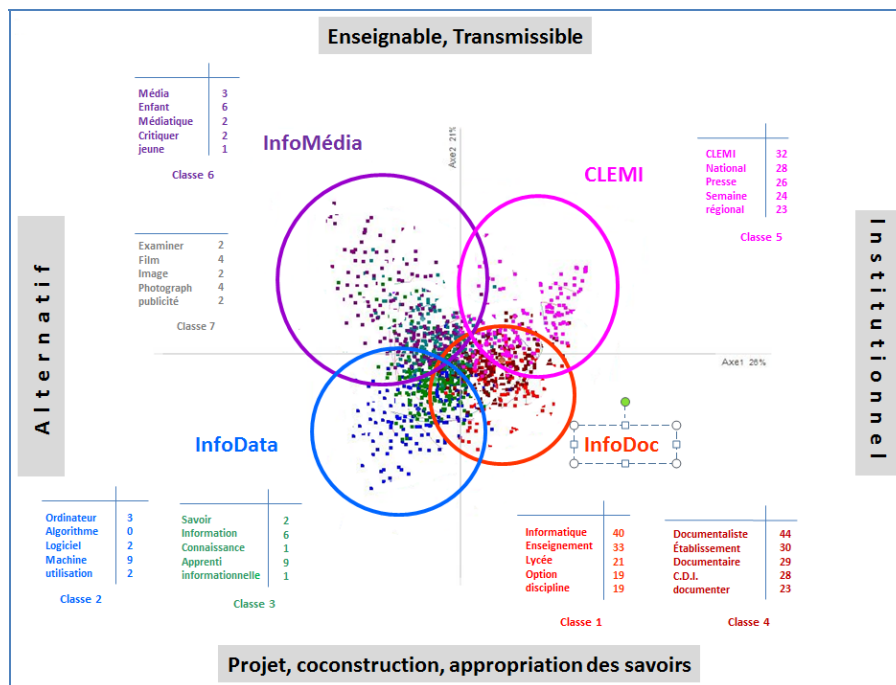


Figure 3. AFC étiquetée

En terme de divergences, on voit plus précisément ce qui distingue les domaines : l'info-doc se situe entre *dispositifs scolaires* d'une part et *enjeux éducatifs* et *contenus enseignables* d'autre part ; l'info-data est concentrée sur les *contenus enseignables* ; l'info-médias se situe plutôt entre *enjeux éducatifs* et *acquisitions par la pratique*.

On retient deux observations essentielles pour notre propos :

- Dans le corpus 1 les discours des 3 domaines apparaissent clairement distincts et ce quel que soit leur lieu d'énonciation.
- La CDH fait apparaître les divergences marquées entre les 3 domaines, en termes d'enjeux éducatifs, de dispositifs scolaires, de contenus à enseigner et de mode d'acquisition des compétences.

3.4.1. Statistiques des occurrences de formes : recherche de convergences

Parmi toutes les observations faites sur le premier corpus, on retiendra ici l'analyse des tableaux statistiques des occurrences de formes. Nous avons exploité ces données pour une approche des convergences discursives entre les domaines. Il s'agissait ici de trouver une méthode permettant de distinguer les mots à caractère *doxique* au regard de l'ensemble du corpus, et ceux à caractère *spécifique* de la classe dans laquelle ils sont positionnés par Alceste.

La méthode consiste à comparer la significativité d'une forme dans sa classe (la valeur du Khi2) avec le volume et la répartition de ses occurrences dans le corpus. Un outil produit par Alceste, parmi les résultats d'analyse du corpus, nous a permis de faire cette observation comparative. Il s'agit du tableau des résultats statistiques sur les occurrences des formes classées (fichier *Excel* intitulé *Présences*). Dans ce tableau, où les formes sont classées par Khi2 dans chaque classe, on trouve notamment les effectifs de ces formes (en nombre d'UCE où elles figurent) dans la classe et dans le corpus. Une colonne donne le calcul du pourcentage du nombre d'UCE de la classe contenant le mot, par rapport au nombre d'UCE du corpus contenant le mot. Nous proposons l'interprétation selon laquelle plus ce pourcentage est élevé, plus le mot est concentré dans la classe ; plus il est faible plus le mot est réparti dans le corpus. Un petit extrait de ce tableau, concernant les termes les plus significatifs de la classe 1, est représenté dans la figure 4.

FORME	KHI2	Nombre d'UCE du Corpus contenant le mot	Nombre d'UCE de la classe contenant le mot	% De la colonne D/Cx100	CLASSEF	Effectif du mot dans la classe	PHIS
informatique	961	796	367	46%	1	559	.41
enseignement	664	634	281	44%	1	367	.34
lycee	295	194	102	53%	1	120	.23
disciplin	269	406	156	38%	1	187	.22
option	236	48	41	85%	1	43	.20
science	207	240	102	43%	1	128	.19
brevet	196	55	41	75%	1	50	.18
secondaire	188	69	46	67%	1	47	.18
college	182	234	96	41%	1	105	.18
primaire	174	78	48	62%	1	50	.17
specialite	160	35	29	83%	1	34	.17
universit	160	142	67	47%	1	75	.17
concours	129	56	35	63%	1	49	.15
introduction	129	70	40	57%	1	41	.15
pays	128	98	49	50%	1	64	.15
enseigner	121	631	165	26%	1	208	.14
niveau	117	279	93	33%	1	110	.14
terminal	112	53	32	60%	1	38	.14
programme	100	294	92	31%	1	106	.13
rentree	99	33	23	70%	1	29	.13
licence	91	23	18	78%	1	20	.12
obligatoire	89	33	22	67%	1	23	.12
professeur	85	262	81	31%	1	81	.12
seconde	85	48	27	56%	1	28	.12
formation	83	655	154	24%	1	193	.12

Figure 4. Extrait du tableau statistique des occurrences de formes

Nous avons recherché des *formes* classées – c’est-à-dire significatives dans au moins une classe – mais dont la présence est, soit *concentrée* au sein des UCE d’une classe, soit *répartie* dans le corpus. Si leur présence est *concentrée* dans une classe, les mots concernés sont *spécifiques* de cette classe. Si leur présence est *répartie* on peut considérer les *mots* concernés, employés dans de nombreuses UCE du corpus, donc dans divers textes, ont un caractère *doxique* dans le corpus. On ne prend pas ici en compte le caractère éventuellement polysémique des mots concernés : ont-ils le même sens dans les divers lieux d’énonciation ? Cet aspect sera analysé dans une phase ultérieure de la recherche (voir § 3.5.2).

Quelle Interprétation peut-on avancer ?

– Aucune forme ne dépasse le pourcentage de 85 % : les mots spécifiques de la classe ne le sont donc globalement pas de façon très marquée.

– Ainsi les formes les plus significatives (Khi2 élevé) « Informatique » et « enseignement », dont les effectifs totaux sont importants, figurent dans les UCE de cette classe pour moins de la moitié de ces effectifs totaux. Elles sont donc plutôt réparties.

– Par exemple, la forme « option » (85 % d’occurrences dans la classe) peut être considérée comme une forme concentrée dans cette classe, le mot est donc bien *spécifique* de cette classe.

– A l’inverse, la forme « formation » qui a un fort effectif dans le corpus, bien que classée en classe 1, n’y figure que pour 24 % de ses occurrences dans le corpus. Le mot « formation » est donc un mot *réparti* dans le corpus et relève probablement d’une doxa partagée.

Nous avons élargi ce relevé des mots *répartis* à l'ensemble des formes classés du corpus, Voici la liste des formes significatives, en termes d'effectifs et de Khi2, dont le pourcentage classe/corpus est inférieur à 30% dans au moins 1 classe :

formation (dans 3 classes), enseigner (dans 2 classes), élève (dans 2 classes), année (dans 2 classes), école (dans 2 classes), utiliser (dans 2 classes), éducation (dans 2 classes), comprendre (dans 2 classes), technolog., général, proposer, proposition, cours, intégrer, associat., national, classe, informatique, mode, discipline, objectif, numérique, évaluation, action, projet, pédagog., établissement, scolaire, mission, moyen, réaliser, travaux, assur., société, considérer, communic., productif, fondamental, citoyen, culture, exprimer, vie, manière, public, offrir, milieu, prendre, différent, demander, étude, langage, média, cas, choix, présenter, type, vue, présenter, aspect, personne., clé, perspective, aborder, approcher, représentation, analyser, journal, effet, égard, présent.

Dans le corpus global, il existe donc des termes significatifs (classés) et plutôt répartis. Cette liste apparaît comme une intersection discursive entre les trois domaines et il faut constater qu'elle ne contient que peu de termes littéraciques ou translittéraciques (on trouve cependant : informatique, numérique, langage, communic., productif, média, vue, culture, représentation, analyser, journal, effet ...) Observée dans sa globalité, il apparaît que cette liste se situe dans le champ sémantique large de *l'éducation scolaire*. L'étiquette qui la caractérise pourrait bien même être celle de la *pédagogie innovante*.

3.4.2. Sous-corpus à domaine unique : les lieux d'énonciation sont les plus clivants

Les sous-corpus du corpus 1, réduits à un domaine unique (info-doc, info-médias et info-data), ont également été analysés. Les tableaux statistiques des occurrences présentent, dans chaque cas, des pourcentages corpus/classes beaucoup plus élevés que dans le corpus 1 global. Dans tous ces sous-corpus le nombre de formes *concentrées* dans une classe (pourcentage au-dessus de 85%) y est très élevé, et très peu de formes présentent un pourcentage inférieur à 30% (ce qui indiquerait des termes *répartis*). Chacune des classes de ces sous-corpus présente donc un vocabulaire assez peu partagé par les autres.

Or nous avons également observé que dans ces sous-corpus de domaine unique, les classes sont très clivées par la variable « Lieu d'énonciation » (institutionnel, professionnelle ou universitaire). Voici les conclusions de ces observations :

Dans le corpus info-doc, où Alceste distingue 3 classes, la classe 1 (étiquetée « Acteurs institutionnels ») inclut distinctement la valeur « institutionnel » de la variable « lieu » ; la classe 2 (étiquetée « Didactisation ») recouvre le lieu « professionnel » ; la classe 3 (étiquetée « Notions ») celle du lieu « universitaire ».

Dans le corpus Info-data, où Alceste distingue 4 classes, la classe 1 (étiquetée « Contenus inférés à la science informatique ») ne souligne aucune valeur de la variable « Lieu » ; dans la classe 2 (étiquetée « Politiques éducatives »), c'est la valeur « institutionnel » de la variable « Lieu » qui est mise en évidence ; dans la classe 3 (étiquetée « Contenus inférés aux pratiques sociales des élèves »), c'est la valeur « universitaire » qui apparaît, et dans la classe 4, c'est la valeur « professionnel » qui se distingue.

Dans le corpus info-médias, où Alceste distingue 6 classes, la classe 3 (étiquetée « Enjeux ») distingue le lieu « universitaire », la classe 1 (étiquetée « Contenus ») distingue le lieu « professionnel », le lieu « universitaire » se caractérise quant à lui dans les classes 4 (« cadres

institutionnels »), 5 (« Formation ») et 6 (« Semaine de la presse et des médias dans l'école ». La classe 2 (« Compétences ») ne distingue pas de lieu particulier.

Au total, il semblerait que les éléments de langage existent plus en fonction des *lieux* (institutionnel, universitaire, professionnel) qu'en fonction des *domaines*. Il apparaît en tous cas que les langues de spécialité se distinguent selon les lieux d'énonciation.

Ces lieux apparaissent comme assez clivants du point de vue lexical. Il importait donc de poursuivre la recherche en se concentrant sur un seul lieu d'énonciation. Nous avons donc choisi de poursuivre notre analyse sur un seul d'entre eux, le lieu d'énonciation institutionnel, celui qui contient les injonctions pour les pratiques enseignantes à venir. C'est ainsi qu'est née l'idée de constituer le nouveau corpus, intitulé « corpus 2 » fait de textes institutionnels récents, notamment ceux de l'année 2015, année fondatrice en matière de textes institutionnels translittéraciques.

3.5. Résultats et interprétations du corpus 2 : vers l'effacement des dissensus ?

Les propositions politiques de transformations (refondation de l'Ecole...) se multiplient, et des textes officiels émanant notamment du Ministère de l'Éducation Nationale incitent à éduquer aux Médias et à l'Information. Des significations socialement attestées et temporairement stabilisées s'affirment. L'objectif principal de l'approche développée avec le corpus 2 consistait à prendre en compte sur la pluralité des conceptions de l'EMI (et en conséquence des « vertus » des 3 domaines), et de mettre l'accent sur un certain ajustement comme élément fondamental de l'élaboration d'un discours partagé. Dans le cadre de notre recherche se posaient et se posent les questions de l'élaboration et la circulation de savoirs dans leur diversité épistémologique (savoirs savants, savoirs professionnels, savoirs traditionnels, savoirs issus de la pratique, savoirs profanes ...) et la recomposition de savoirs disciplinaires. Car, comme l'écrit Edgar Morin, « les concepts voyagent et il vaut mieux qu'ils voyagent en sachant qu'ils voyagent » (1990).

3.5.1. Cartographie du corpus 2 : les domaines bousculés

La technique de la CDH donne une arborescence de 5 classes (figure 5).

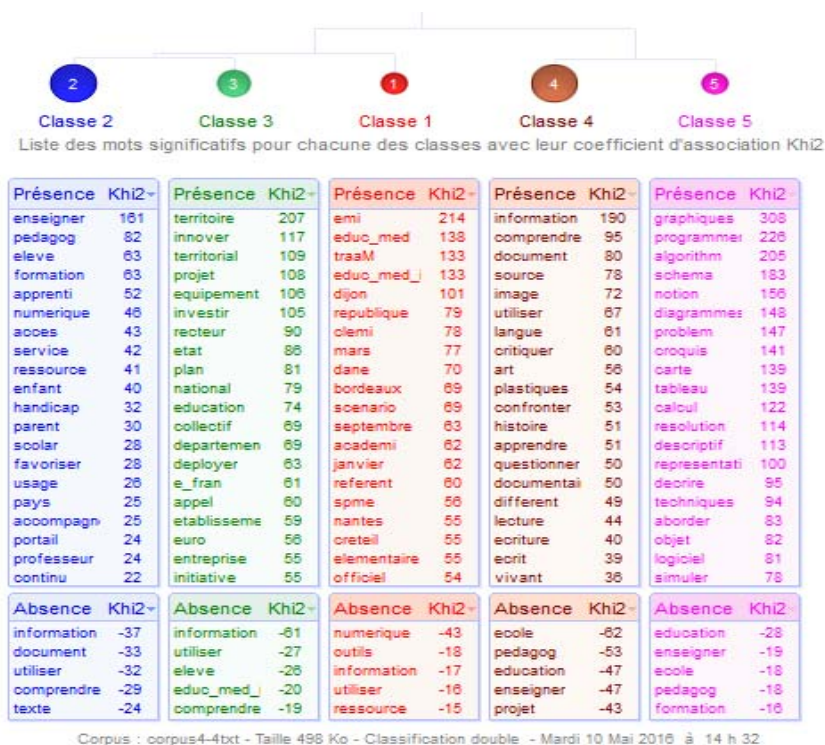


Figure 5. Le tableau des présences/absences de la CDH, définissant les classes du corpus 2

La classe 4 est la plus volumineuse, elle concentre 32% des UCE. Les classes 1 et 5 sont les plus petites, avec 12% des UCE chacune.

L'étiquetage de l'arborescence des classes et des groupes de classes donne le diagramme d'organisation suivant (figure 6).

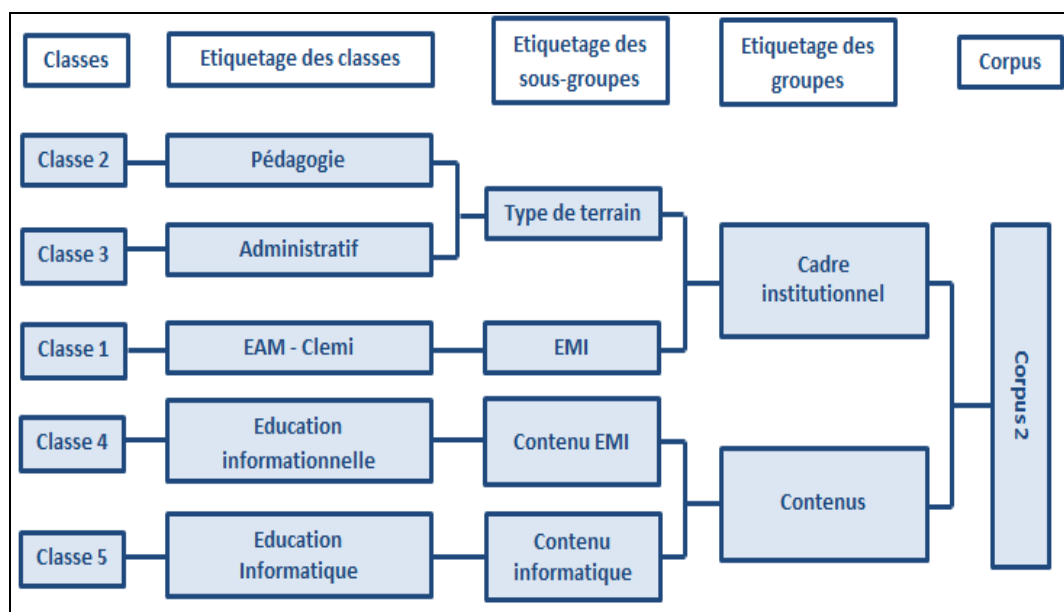


Figure 6. Dendrogramme des univers sémantiques du corpus 2

Interprétation

- Il s'agit bien ici d'un corpus exclusivement institutionnel de textes injonctifs, donc unifié du point de vue de l'énonciation, dans lequel on retrouve nettement les domaines du corpus 1 : info-médias, info-doc et info-data. Mais ils se situent dans une distribution très différenciée : la petite classe 1, qui évoque l'info-média par ses structures autour des actions du Clemi, se retrouve dans le groupe « cadre institutionnel ». Les classes 4 et 5 qui évoquent l'info-doc et l'info-data sont distinguées par leur contenu, mais avec des volumes très inégaux.

- La distinction entre les 3 domaines d'origine devient une double distinction binaire entre *contenus* et *cadre institutionnel* d'une part, *éducation informationnelle (info-doc)* et *informatique* d'autre part.

- On remarque dans la classe 2 une absence des termes liés à la dimension technologique de l'informatique (hardware...), au profit des termes liés à la dimension algorithmique dominante.

Conclusion

Cette distribution nouvelle semble effacer les *dissensus*, en particulier entre les domaines info-média et info-doc. Mais il convient d'y regarder de plus près. En 2015 en effet (date de la presque totalité du corpus2), dans le contexte des attentats, on évoquait fréquemment une nécessaire éducation aux médias dans les prises de position publiques médiatisées. Or l'analyse par Alceste du corpus – fait de textes plutôt injonctifs institutionnels - montre que l'éducation aux médias apparaît surtout dans les questions de structures institutionnelles, du fait de l'existence du Clemi, et pratiquement pas dans le contenu pédagogique. Quant à la classe 4, qui devrait intégrer l'environnement sémantique des contenus l'EMI, elle évoque surtout celui de l'info-doc. Ce décalage interroge la validation (au sens de G-E Sarfati) de l'étiquette « EMI ». Par ailleurs le domaine info-data reste très isolé des autres et, semble-t-il, incomplètement représenté. La nouvelle construction doxique serait-elle véritablement légitime dans ses références aux 3 cultures ?

3.5.2. *Éléments de légitimation dans la nouvelle construction doxique*

Nous avons donc tenté une recherche plus affinée en confrontant les proximités de *formes* relevées par Alceste avec les *termes* - non-polysémiques - retenus par les experts des 3 domaines. Nous esquissons-là un travail sur la sémantique lexicale en comparant un corpus de termes experts et les résultats d'Alceste sur un corpus de textes.

Nous avons choisi cette fois-ci de partir des savoirs experts, c'est à dire de *termes* issus de la parole des experts, dans chacun des 3 domaines, puis de regarder avec Alceste comment se comportent les formes correspondantes dans les textes du corpus 2. Voici la méthode retenue, en 4 étapes :

- Déterminer 10 *termes* dans chaque domaine et les accompagner d'indications sur l'environnement sémantique attendu, correspondant au sens que lui donnent les experts du domaine ;
- Rechercher des occurrences des *formes* correspondantes dans le corpus ; construire des *réseaux de forme* (Alceste) dans la classe et dans le corpus ;
- Comparer le sens des *mots* avec l'aide des *concordanciers* (Alceste).
- Recherche d'un résultat : Les *formes* analysées sont-elles significatives dans le sens proposé par les experts ? Les *termes* des experts sont-ils effectivement représentés par les formes relevées dans le corpus ?

Quelques obstacles induits par cette méthode

- Rares sont les termes des experts qui présentent un *réseau de forme* dans Alceste. En effet certaines formes présentent des occurrences nombreuses mais qui ne sont pas assez significatives pour qu'Alceste dégage un *réseau de formes*.
- Les formes analysées peuvent correspondre à des mots polysémiques et pas nécessairement utilisés dans le sens des termes experts. Il faut analyser les *concordanciers de formes*, ligne par ligne pour en distinguer le sens.

Mise en oeuvre

Afin de déterminer une liste de 10 termes reconnus par les experts de chaque domaine, nous avons soumis une enquête à des experts du domaine. Ainsi pour l'éducation aux médias, une liste de 18 termes-concepts-notions, leur a été présentée comme :

- traditionnels de l'EAM,
- légitimant la singularité du domaine EAM,
- assurant la viabilité du discours de ce domaine dans le long terme.

Il a été demandé aux experts de :

1. Cocher les termes qui vous paraissent répondre à ces critères.
2. Barrer ceux qui, selon vous, n'y répondent pas.
3. Ajouter éventuellement jusqu'à 3 autres termes.

Une nouvelle liste a ainsi été établie en tenant compte des réponses des experts, de façon à conserver 10 termes principaux, et pour chacun des termes un environnement sémantique attendu au regard du sens recherché. Une exploration de même nature a été réalisée dans les 3 domaines, en mettant à contribution une dizaine d'experts dans chacun.

Observations des réseaux de formes et des concordanciers

En utilisant les résultats Alceste du corpus 2, il s'agissait ensuite de comparer les formes, dans le corpus, de chaque terme expert, son ou ses éventuels *réseaux de formes*, avec les termes attendus par les experts et leur environnement sémantique. L'existence d'un éventuel synonyme a été prise en compte. Nous avons également observé les *concordanciers des formes* pour mieux évaluer le sens dans lequel chaque mot apparaît. Les concordanciers permettent en effet de visualiser le contexte d'utilisation d'une forme en affichant x mots avant et x mots après chaque occurrence dans une UCE.

3.5.3. Interprétation

Les observations sont différentes selon les domaines.

Pour l'éducation aux médias (info-média)

Les mots retenus comme termes des experts sont les suivants :

citoyen (<i>empowerment</i> - autonomisation - responsabilité - engagement – débat).
message (transmission-circulation-transformation-déformation de contenus médiatiques)
critique (esprit critique - pensée critique vis à vis des médias)
politique (éducation au politique, au vivre ensemble)
créativité (productions médiatiques des élèves-scolaires-jeunes)
publication (comme acte social, cadre légal)
analyse (des objets-discours-textes-sources médiatiques)
presse (tous les supports de l'info-news, source, journalisme, entreprise de presse)
audiovisuel (image, son, vidéo, film, radio, télévision, multimédia, jeux)
Internet (comme phénomène de société, mais aussi nouvelle économie de l'information)
culture (médiatique, représentations)

Les réseaux de formes n'existent que pour 4 des termes proposés et, à chaque fois, uniquement dans la classe où la forme correspondante est classée.

En observant avec Alceste le comportement dans le corpus 2 des termes donnés par les experts, on voit que les termes principaux sont présents, mais que les mots (ou des synonymes) attendus dans leur environnement sémantique ne figurent pas dans les réseaux de formes éventuels. En d'autres termes, on peut dire que des *mots* sont présents dans le corpus, mais souvent avec un sens différent de celui donné par les experts. Et donc qu'il ne s'agit pas de leurs *termes*. La doxa des experts du domaine info-médias semble donc mal représentée dans le corpus 2.

Pour l'éducation à l'informatique (info-data)

Les mots retenus comme termes des experts sont les suivants :

algorithme

ordinateur (machine, unité centrale, disque dur, périphérique, processeur)

réseau

programmation (langage, programme, programmer)

mémoire (ou mémoire vive)

données (base de données, data)

logiciel

fichier

binaire

bureautique

On constate une représentation encore plus faible des notions fondamentales, selon les experts de l'info-data, dans les textes du corpus 2. Exemples :

- Parmi ces termes, les 3 formes significatives, les plus représentés sont : « algorithme » (18 occurrences), « programmer » (11 occurrences) et « logiciel » (63 occurrences) .Seules ces 3 formes présentent un réseau de forme. Ils sont tous dans la classe 5 et pas du tout répartis dans le corpus.

- Le mot « logiciel », est présent le plus souvent au sens de l'utilisation d'applications diverses, mais il n'a pas un Khi2 élevé (82 dans la classe 5).

- Le mot « donnée » est présent 93 fois, mais sans réseaux de formes car le terme est réparti (avec un Khi2 faible), et rarement dans le sens de « donnée informatique » ;

On conclut à une faiblesse de la présence des mots de l'informatique et quasi-absence des termes des experts. L'info-data est un langage de spécialité qui ne percole pas dans les autres spécialités, et la doxa des experts de l'info-data semble donc pratiquement absente du corpus.

Pour l'éducation à l'information (info-doc)

Les mots retenus comme termes des experts sont les suivants :

archive (archivage, mémoire, trace)

auteur (autorité, responsabilité, identification, droit)

crédibilité (fiabilité, information, source)

document (ressource, production, indexation, publication, support, auteur, documentarisation, collecte)

indexation (mot-clé, tag, classification, langage documentaire, thésaurus, descripteur, vedette-matière, automatisée...)

information (droit, espace, évaluation, recherche, source, besoin, traitement, technologie de, système, classement, collecte)

média (médiation, support, technique, source, communication, message, presse)

pertinence (de la requête, bruit, silence ; de l'information ; évaluation de la pertinence)

Recherche (moteur, référencement, d'information)

Veille (informationnelle, stratégique, pédagogique ; outil, curation)

Les termes proposés par les experts sont beaucoup plus présents en nombre d'occurrences de formes. Seuls quelques termes très spécifiques du domaine, comme « indexation » (4 occurrences) manquent à l'appel dans le corpus. On constate nettement une plus forte présence des mots de l'infodoc. Le nombre d'occurrences de la plupart d'entre eux est nettement plus élevé : « information » au sens de l'infodoc, présente 430 occurrences, « document », en présente 139. La plupart de ces occurrences correspondent au sens attendu par les experts, ce qui est attesté par l'observation des concordanciers. L'analyse des 6 réseaux de formes existants montre la forte cohérence de ces formes dans le corpus : « Information » et « document » présentent un réseau de formes dans le corpus complet et un autre dans la classe 4 ; « recherche » a un réseau de formes dans le corpus complet (avec « information » comme forme la plus proche, donc au sens de « recherche d'information »), et « auteur » présente un réseau de forme dans la classe 4. De plus, la plupart des autres termes – y compris ceux cités dans l'environnement sémantique - figure en bonne place dans ces réseaux de formes. La langue de spécialité du domaine semble exister largement dans ce corpus.

Discussion

Certes les listes de termes sélectionnés ne constituent pas, loin de là, un lexique des langues de spécialité des trois cultures. Des biais méthodologiques peuvent bien sûr être discutés, notamment quant au choix des experts, leur nombre, l'influence de la liste proposée, etc.

De plus le nombre réduit de termes ne permet pas de définir, à cette étape de la recherche, une délimitation ou une légitimation complète des discours. Cependant nous avons considéré que la méthodologie de constitution de ces listes leur confère une qualité d'échantillon significatif qui permet de repérer les tendances d'intégration de ces langues de spécialité, de dilution ou de domination doxique.

3.5.4. Esquisse d'une organisation de la doxa

En s'appuyant sur le modèle (G-E Sarfati, 2002) d'une organisation effective du "territoire imaginaire qui fonde l'univers de croyance", on peut chercher ici des éléments de réponses aux questions suivantes : comment s'organisent les discours narratifs, s'appuyant sur le passé, définissant le présent et projetant un avenir ? Peut-on identifier les traces des 3 cultures dans les discours institutionnels récents ? Voici les éléments de réponses que cette recherche permet de formuler :

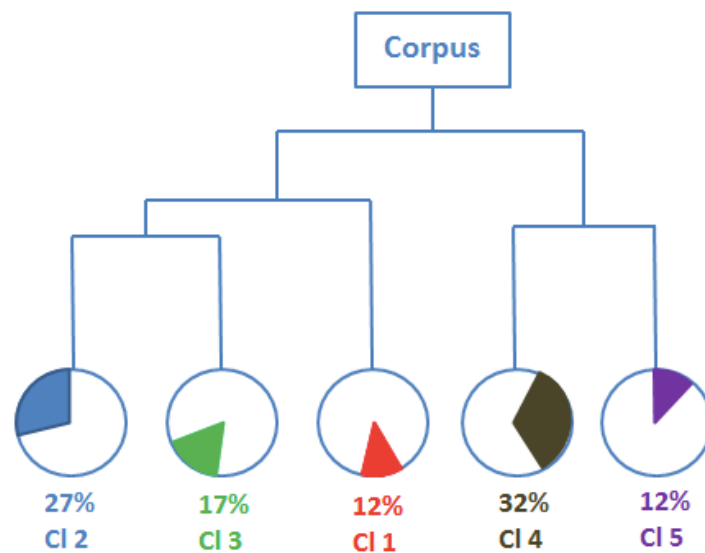


Figure 7. Pourcentages d'UCE qui définissent chaque classe du corpus 2

Différenciation de la culture informatique : on ne retrouve, dans le corpus 2, ni les *mots* de la culture informatique ni les *termes* des experts.(12 % des UCE sont suffisants pour définir la classe 5).

Limitation de la culture info-médias : on retrouve les *mots* de la culture info-médias mais pas les *termes* des experts.(12 % des UCE sont également suffisants pour la définir la classe 1).

Pérennisation de la culture Infodoc : on retrouve les *mots* de la culture info-doc comme les *termes* des experts.(32 % des UCE sont nécessaires pour définir la classe 4). Alceste a besoin de plus d'UCE (segments de textes) pour retrouver une spécificité, Le discours de l'infodoc est le plus "neutre", plus dilué, moins spécifique au plan de l'analyse statistique factorielle.

Cette approche conduit à une forme d'interprétation du présent à la lumière du passé. Force est de constater une réelle percolation des termes issus de l'info-doc dans l'ensemble du corpus 2. Peut-on conclure de ce fait à une "domination" de la doxa info-doc ? Ce serait en effet le cas si l'on observait une absorption effective des deux autres cultures. Mais il semble bien qu'il n'en soit rien : les cultures de l'info-média, et surtout l'info-data, pourraient bien avoir tendance à être occultées dans les approches translittéraciques.

De ce fait une question apparaît légitimement : quel est l'avenir pour les « savoirs experts » de l'info-média et de l'info-data dont l'essence historique semble gommée dans ce corpus de textes récents d'injonctions et de recommandations institutionnelles ?

Conclusion

En quoi la convergence de trois domaines et son agencement en doxa peut-il intéresser le chercheur, le politique et le praticien –impliqués ou pas- et constituer pour eux un objet pertinent d'enquête, d'analyse et de réflexion ? Il est vrai, que sous le regard des chercheurs, la convergence des différentes cultures de l'information et l'EMI cessent d'être perçus maintenant comme une composante nouvelle et évidente du monde scolaire, mais apparaît bien plutôt comme un objet « socialement construit », comme le produit d'un processus permanent d'élaboration et d'institutionnalisation dans lequel se traduisent à la fois des mobilisations de savoirs disciplinaires et de ressources, des confrontations d'intérêts, des conflits de valeurs et des enjeux de pouvoir.

Le système éducatif français ne se borne pas à sélectionner parmi les savoirs et des pratiques de références, elle doit aussi les organiser, les livrer à un travail d'agencement et de mise en mots à fins

politiques et éducatives. Ces « arbitrages » fournissent une grille de lecture intéressante pour saisir la fabrication d'une nouvelle convention (ce que nous examinons avec le concept de doxa). C'est-à-dire un ensemble de critères explicites et implicites, un ensemble de repères auxquels la communauté éducative peut adhérer et les opérations de mise en forme et de mise en récit qu'implique sa mise en oeuvre. C'est donc à une compréhension des processus de qualification que nous nous sommes confrontés.

Mais il faut aller plus loin, dans l'analyse des savoirs que l'école se doit d'enseigner et dessiner les contours d'une « transposition didactique » de l'EMI dont les fondements scientifiques sont sérieusement ancrés dans des disciplines diverses, et dont les pratiques sociales dites « de référence » sont en mutation constante. C'est bien cette situation particulière à la fois de diversité et de labilité qui constitue la spécificité d'une éducation à la culture numérique et en particulier de l'Education aux Médias et l'Information. Sans doute conviendra-t-il ainsi, dans le cadre d'une approche comparative, de différencier les discours scientifiques, professionnels et institutionnels dans l'espace francophone. Il sera nécessaire d'étudier ainsi plus en détail le poids paradoxal de l'« info-doc » dans d'autres textes institutionnels. Peut-on faire l'hypothèse, qu'en France, la production discursive des acteurs de l'info-doc a fini par imprégner l'ensemble des textes, en raison du fait que l'info-doc est le seul des trois domaines à pouvoir s'appuyer sur une profession instituée ? *A contrario*, est-ce un hasard si les recherches sur l'éveil à l'information au niveau de l'enseignement élémentaire ne sont pas examinées ? Il conviendra de situer alors comment la volonté politique s'ordonne en fonction de la diffusion scientifique et professionnelle et explique des différences.

La construction des « objets d'enseignement », adossée aux prescriptions institutionnelles, doit en effet se négocier dans une interaction constante entre les questions posées à l'école par la société (et par sa jeunesse), les questions que l'école se pose à elle-même, et la vigilance des disciplines scientifiques. Les différentes composantes de l'EMI n'ont certes pas vocation à transposer directement ces disciplines scientifiques, mais c'est bien en termes de savoirs disciplinaires que se légitime la réflexion sur le formatage scolaire (ce que l'on a évoqué avec la notion de grammaire). Selon nous, à un premier niveau s'opposent deux visions du développement des connaissances, l'une portée par les tenants de la Science de l'Informatique, l'autre par les Sciences de l'Information et de la Communication. A un second niveau, la chose est moins nette, on note la volonté de changer à la fois le contenu des enseignements -les sacro saintes disciplines scolaires- et la manière de le transmettre. Certains résultats montrent que pointe l'enseignement par projet, les pratiques de questionnement du monde, et l'ancrage dans les Sciences de l'Education. S'agit-il pour autant d'une perspective a-disciplinaire ou pluridisciplinaire et unifiante ?

Bibliographie

- [LEM 14] LEMAIRE M., *Mécanique et incertain*, ISTE Editions, Londres, 2014.
- [CHE 13] CHETTO M., *Ordonnement dans les systèmes temps réel*, ISTE Editions, Londres, 2013.
- [COM 03] COMPTE C., « Enjeux de la formation ouverte et à distance : technologies et apprentissages », dans I. SALEH, D. LEPAGE et S. BOUYAHI (dir.), *Les TIC au cœur de l'enseignement à distance*, Actes Huit Coll., Université de Paris VIII, 2003.
- [BEL 94] BELARBI A., HSU T.-T.-C., « Constitutive laws of concrete in tension and reinforcing bars stiffened by concrete », *ACI Structural Journal*, titre 91-S46, p. 465-474, juillet/août 1994. [BEL 96] BELISLE C., LINARD M., « Quelles nouvelles compétences des acteurs de la formation dans le contexte des TIC ? », *Education Permanente*, n° 127, p. 19-47, 1996.
- [BEL 96] BELISLE C., LINARD M., « Quelles nouvelles compétences des acteurs de la formation dans le contexte des TIC ? », *Education Permanente*, n° 127, p. 19-47, 1996.
- [COM 03] COMPTE C., « Enjeux de la formation ouverte et à distance : technologies et apprentissages », dans I. SALEH, D. LEPAGE et S. BOUYAHI (dir.), *Les TIC au cœur de l'enseignement à distance*, Actes Huit Coll., Université de Paris VIII, 2003.

- [ADA 11] ADAM J.-M., *Les textes : types et prototypes*, (3e édition revue) Armand Colin, Paris, 2011.
- [AMO 99] AMOSSY R., (éd) *Images de soi dans le discours. La construction de l'ethos*, Delachaux et Niestlé, Lausanne, 1999
- [AMO 10] AMOSSY R., *La présentation de soi. Ethos et identité verbale*, PUF, Paris, 2010.
- [AMO 14] AMOSSY R., *Apologie de la polémique*, PUF, Paris, 2014.
- [BRO 96] BRONCKART J.-P., *Activité langagière, textes et discours. Pour un interactionnisme socio-discursif*, Delachaux et Niestlé, Lausanne, 1996.
- [DEL 13] DELAMOTTE E., « La culture informationnelle : le noyau central d'une nouvelle forme scolaire ? », dans PAPY F., (dir.) *Recherches ouvertes sur le numérique*, Cahiers du numérique, Hermès Lavoisier, Paris, p. 263-281, 2013.
- [DEL 17] DELAMOTTE E., JEANNE P., CHENEVEZ O., LOICQ M., SERRES A., FLUCKIGER C., BALLARINI-SANTONOCITO I., COMPANT la FONTAINE M.-L., « Convergences et divergences des trois cultures de l'information (Info-doc, Info-média, Info-data) : deux corpus de textes sous la loupe d'Alceste », Rapport de recherche de la Tâche 1 de l'ANR Translit, mars 2017. Prochainement disponible sur HAL.
- [DOU 11] DOUEIHI M., *Pour un humanisme numérique*, Seuil, Paris, 2011.
- [LEM 94] LE MAREC J., « Situations de communication dans la pratique de recherche : du terrain aux composites », *Études de communication*, n° 25, [En ligne]: <http://edc.revues.org/index831.html>
- [MAR 98] MARCHAND P., *L'analyse du discours assistée par ordinateur*, Armand Colin, Paris, 1998.
- [MOR 90] MORIN E., *Introduction à la pensée complexe*, Paris, Seuil, 1990.
- [ORI 04] ORIGGI G., « Croyance, déférence et témoignage », dans PACHERIE E., PROUST J., (éd) *La philosophie cognitive*, Ophrys/MSH, Paris, 2004.
- [ORI 08] ORIGGI G., *Qu'est-ce que la confiance ?* Vrin, Paris, 2008
- [REI 90] REINERT M. « Alceste une méthodologie d'analyse des données textuelles et une application : Aurélia de Gérard De Nerval », *Bulletin de méthodologie sociologique*, 26, p. 24-54, 1990.
- [REI 92] REINERT M., *Manuel d'utilisation du logiciel ALCESTE*, Inédit, 1992.
- [REI 93] REINERT M., « Les « mondes lexicaux » et leur « logique » à travers l'analyse statistique d'un corpus de récits de cauchemars », *Langage et société*, 60, p. 5-39, 1993.
- [SAR 02] SARFATI G.-E., « Qu'est-ce qu'un texte canonique », dans POFFET J.M., (dir). *L'autorité de l'écriture*. CERF, Paris, 2002.
- [THO 74] THOM R., *Modèle mathématique de la morphogénèse*, Paris, Presses de la Cité, 1974.