

Rubrique « Vie de la recherche en formation des adultes » commune aux revues *Savoirs* et *TransFormations*

N° 4 – Publications de recherche concernant la formation des adultes : comparaison et combinaison des méthodes de repérage (novembre 2017)

Olivier Las Vergnas, CIREL-Trigone (EA 4354) et CREF-AFA (EA 1589).

Résumé :

Cet article est le quatrième d'une rubrique dédiée à la vie de la recherche francophone concernant la formation des adultes. En l'absence d'une définition partagée de ce champ et d'un système d'indexation structuré permettant de l'explorer, les trois premiers articles ont été consacrés à des investigations complémentaires entre elles, visant à fournir plusieurs visions triangulaires de la production écrite, utilisables par la suite pour en suivre en particulier l'évolution temporelle. Ainsi, deux articles ont été focalisés sur les thèses, s'appuyant tous les deux sur la base nationale Sudoc. Un troisième s'est intéressé à des moyens d'identifier toutes les publications de recherche déposées dans l'archive ouverte Hal-SHS soit via une requête unique sur les mots « formation ou éducation des adultes », soit en juxtaposant une série de requêtes sur des mots beaucoup plus précis, correspondant à une exploration par facettes. En considérant que la production à prendre en compte est constituée de l'ensemble des documents recueillis par l'une ou l'autre des méthodes, le champ peut se regarder comme composé de cinq classes. Deux sont à l'échelle macroscopique (emploi/entreprise et histoire sociale), une autre se focalise sur les TICE et les pédagogies liées, et les deux dernières très imbriquées correspondent à la didactique, aux apprentissages et à la professionnalité, liées aux professions de santé.

Abstract:

This article is the fourth in a section devoted to the life of Francophone research on adult education and training. In the absence of a shared definition of this field and a structured indexing system to explore it, the first three articles were devoted to complementary investigations, aiming at providing several triangular visions of written production. Two of them were focused on PhD theses, both based on the Sudoc national database. A third looked at ways to identify all research publications in the Hal-SHS open archive either through a single query on the words "training or adult education" or by juxtaposing a series of queries on more precise words, corresponding to an exploration by facets. Considering that the production to be taken into account consists of all the documents collected by one or other of this three methods, this field is seen as composed of five classes. Two are on the macroscopic scale (employment / enterprise and social history), another focuses on ICT and related pedagogies and the last two on "didactics", "learning" and "professionalism", linked with health professions.

Mots clés : publications de recherche, bibliométrie, lexicométrie, formation des adultes, analyse lexicale

Keywords: research publications, bibliometrics, lexicometry, adult education, lexical analysis

Note de l'auteur : les dernières données rapportées ici correspondent à des requêtes effectuées dans HAL-SHS fin février 2017.

1. Rappel des travaux précédents : trois visions de la définition de « formation des adultes »

Cet article est le quatrième de la nouvelle formule¹ de la rubrique « Vie de la recherche en formation des adultes » ; il s'inscrit dans la volonté globale de proposer un état dynamique des lieux de la production de connaissances dans ce champ, au moins dans l'univers francophone pour le moment.

Les deux premiers articles ont été fondés sur l'analyse des thèses répertoriées dans la base de données officielle du Service universitaire de documentation (Sudoc) et le troisième sur celle des documents déposés dans l'archive ouverte Hal-SHS (Las Vergnas, 2016a, 2016b et 2017). Cette base est en effet devenue aujourd'hui l'archive ouverte de référence des publications en SHS francophones, regroupant quant à elle non seulement des thèses, mais aussi des communications scientifiques plus collectives et/ou courtes.

Ces trois premiers articles de cette rubrique ont donc proposé trois méthodes différentes susceptibles de fournir une vision de la production de recherche en formation des adultes.

La première (M1, publiée dans le n° 1) a consisté à rechercher les thèses d'université (excluant donc les thèses dites d'exercice, en médecine, pharmacie ou de sage-femme) répertoriées dans le Sudoc contenant les termes génériques que sont « formation des adultes », « éducation des adultes », « formation tout au long de la vie » en mots clés ou dans leur résumé, aboutissant à 41 thèses en 2014-2015 puis à 275 thèses dans la totalité du Sudoc contenant ces mêmes mots comme mots clés (méthode M1bis publiée dans le n° 2, s'appuyant précisément sur une recherche de ces termes génériques dans les « mots sujets » utilisés pour indexer dans le Sudoc).

Une deuxième méthode (M2) a été de chercher avec ces mêmes mots génériques (en mots clés ou dans les résumés) parmi toutes les publications déposées dans Hal-SHS, ce qui a conduit à 172 documents, après filtrage de ceux qui employaient aussi un terme renvoyant à la formation initiale (plus précisément, ont été rejetés tous les résumés contenant « enseignant » ou « élève »).

Enfin, la troisième méthode (M3) a consisté à cumuler pas à pas les résultats d'une batterie de requêtes testant chacune la présence d'un ou plusieurs termes caractéristiques des divers aspects de la formation des adultes (toujours avec le même filtrage des références à « enseignant » ou « élève ». En l'occurrence, ces termes étaient issus du sommaire d'un ouvrage pouvant être considéré comme un référentiel du champ, à savoir le *Traité des sciences et techniques de la formation* de Carré et Caspar, troisième édition, 2011). Cette méthode qui a finalement cumulé près d'une quarantaine de sous-requêtes très thématiques a permis de repérer toujours dans HAL-SAS 1 182 documents différents, dont seuls 55 étaient communs à la liste de 172 obtenus précédemment par M2.

Le tableau T1 récapitule ces différentes méthodes.

Les procédures, difficultés et résultats liés à chacune de ces méthodes ont été abondamment détaillés dans les numéros précédents de cette rubrique. Comme prévu à la fin du n° 3, ce que nous proposons dans ce n° 4, c'est de regarder ce que la comparaison de ces analyses conduites selon des perspectives différentes nous apporte, tant en termes de meilleure compréhension de la production de recherche en elle-même que d'écart de représentation de sa définition selon les méthodes. Dans la suite de cet article, nous emploierons chacun des sigles M1bis, M2 ou M3 aussi bien pour désigner la requête effectuée dans les bases de données que pour désigner les corpus de résumés des textes trouvés par cette requête.

¹ En accord entre les rédactions des deux revues, cette rubrique est commune à la revue *Savoirs* et à la revue *TransFormations*.

| | Bases de données explorées | Requête de recherche | Nombre de publications trouvées | N° de rubrique de publication | Périmètre (pour l'indexation) | Filtrage supplémentaire |
|--------------|--|---|---------------------------------------|-------------------------------|--|-------------------------------|
| M1 | Thèses dans le Sudoc indiquées soutenues en 2014 ou 2015 | « formation des adultes », « éducation des adultes », « formation continue » dans les mots sujets et résumés | 41 thèses | N° 1 | Langage Rameau | Par lecture flottante humaine |
| M1bis | Toutes les thèses dans le Sudoc | « formation des adultes », « éducation des adultes », « formation continue » dans les mots sujets | 275 thèses | N° 2 | Langage Rameau | non |
| M2 | Tous les documents dans Hal-SHS | « formation des adultes », « éducation des adultes », « formation tout au long de la vie », « éducation permanente » dans les mots clés et résumé | 172 documents dont 55 communs à M3 | N° 3 | Auto-indexation | Filtrage de « enseignannt* » |
| M3 | Tous les documents dans Hal-SHS | Batterie de 40 sous-requêtes cherchant un par un les thèmes présents dans le sommaire du « Traité des sciences et techniques de la formation » dans les mots clés et résumé | 1182 documents (dont 55 communs à M2) | N° 3 | Titres des chapitres du traité (Carré et Caspar, 2011) | Filtrage de « enseignannt* » |

Tableau T1. Caractéristiques des 4 méthodes mises en œuvre dans des numéros précédents de la rubrique « Vie de la recherche en formation des adultes »

Cette juxtaposition de plusieurs méthodes de recherche qu'impose cette exploration confronte à deux manques qui empêchent de s'appuyer sur une requête unique. Le premier est le manque d'une représentation consensuelle des limites de la recherche en formation des adultes (par exemple la question se pose de savoir si la formation des enseignants, le *knowledge management* en entreprise ou encore la GPECT en font partie ou non). Le second est le manque d'un système cohérent (un thésaurus ou une ontologie) qui permettrait d'indexer les travaux de manière hiérarchisée afin d'éviter en particulier les indexations trop fines (les travaux sur le développement professionnel des infirmiers ne se repèrent pas naturellement comme un sous-ensemble de la formation des adultes).

Concrètement, cela signifie qu'il n'y aurait pas de sens à chercher entre M1, M2 ou M3 une méthode meilleure qu'une autre, mais que nous sommes confrontés à la possibilité d'englober différents périmètres plus ou moins thématiques quand nous décidons de suivre l'évolution de la vie de la recherche en formation des adultes. Il s'agirait plutôt *a contrario* de voir comment les assembler. Cette limite ne supprime en rien l'intérêt de ce travail, mais cela conduit à considérer qu'il sera surtout pertinent en permettant des comparaisons d'une période à l'autre. C'est d'ailleurs ce que nous proposerons dans les prochains numéros de cette rubrique, comme nous l'avons déjà fait entre le n° 1 et le n° 2 pour la comparaison des thèses entre M1 (2014-2015) et M1bis (thèses présentes dans le Sudoc, quelle que soit l'année de soutenance).

2. Complémentarité entre les diverses méthodes et corpus

Néanmoins, si l'on veut pour la suite pouvoir considérer la recherche en formation des adultes comme un champ stabilisé, une question clef se pose, celle du degré de différence ou de similitude entre les visions apportées par ces diverses méthodes M1bis, M2 ou M3. Le fait d'utiliser soit une définition liée

à la production des thèses (méthode M1bis), soit une définition restreinte à quelques mots clés spécifiques « formation des adultes » (méthode M2), soit enfin une troisième fondée sur le cumul de toute une série de thématiques appartenant au champ (méthode M3) conduit-il à des périmètres très différents ou au contraire se complétant de manière pertinente ? Pour répondre à cette question, nous avons d'abord comparé les trois corpus correspondant respectivement à ces trois définitions grâce aux analyses lexicales effectuées dans les articles précédents via le logiciel Iramuteq (Ratinaud et Dejean, 2009).

Le tableau T2 rappelle – ligne par ligne – pour chaque méthode (M1, M1bis, M2 et M3) les caractéristiques des corpus identifiés ainsi que les découpages en classes lexicales qui en avaient été obtenus avec ce logiciel. Dans ce tableau, les colonnes ont été agencées de manière à faire apparaître les points communs et les différences entre ces diverses partitions.

2.1. Deux classes lexicales toujours présentes, l'une macro et l'autre micro

On constate ainsi que la classe lexicale 1 « Emploi, socio-économie, entreprises, secteurs » traitant des aspects « macro » de la formation apparaît comme une classe à part entière dans tous les corpus, même si sa déclinaison dans les 41 thèses trouvées par M1 semble plus orientée vers le « transfert de technologie » que vers le « marché de l'emploi ». On retrouve aussi que la classe lexicale 2, elle plus « micro » et correspondant à « sujet, identité, reconnaissance, VAE », semble à peu près présente dans tous les corpus, même si elle se décline particulièrement dans la période récente dans le champ professionnel de la santé, vers les infirmiers notamment.

2.2. Trois classes plus spécifiques : apprentissages, TICE, comparaisons historiques ou internationales

La colonne suivante correspond quant à elle aux « apprentissages » des adultes et aux « apprenants », en particulier en lien avec les apprentissages « linguistiques ». Cette même classe lexicale « apprentissage et pédagogique » se retrouve aussi dans les corpus liés à toutes les requêtes, à l'exception de M2, c'est-à-dire que cette approche « micro » n'est pas repérable en tant que telle dans les documents indexés explicitement comme « formation des adultes » dans Hal-SHS.

La colonne suivante du tableau permet ensuite de constater que la classe 4 « technologies éducatives » ou TICE est repérée dans Hal-SHS (en M2 et M3) mais peu dans le corpus des thèses du Sudoc (M1 et M1bis) ; on observe donc ainsi qu'alors qu'il y a dans Hal-SHS de nombreux articles portant sur les TICE mobilisées spécifiquement au service de la formation des adultes, ce sujet n'est que peu signalé dans les mots sujets des notices des thèses du Sudoc.

Enfin, la classe lexicale 5 concerne les études de socio-économie comparative à dimension historique (ou internationale), qui sont peu présentes dans les thèses récentes. Cela correspond à ce que nous avons remarqué dans les articles précédents, à savoir que ce sujet n'est plus aujourd'hui traité ou indexé comme entrant dans le cœur de recherche en formation des adultes. Il est néanmoins à noter – comme nous le verrons plus loin – que cette classe lexicale reste existante si l'on considère le sous-ensemble des documents trouvés par M2 et non simultanément trouvé par M3 (c'est-à-dire parlant de comparaison historique sans se focaliser en même temps sur une sous-thématique clef de la formation des adultes).

| Méthode de requête utilisée | Nbre de publications du corpus obtenu | Classe lexicale 1 = Emploi, socio-économie, entreprises, secteurs | Classe lexicale 2 = Sujet, Identité, Reconnaissance, VAE | | Classe lexicale 3 = Apprentissage des adultes, (en particulier langage) | Classe lexicale 4 = TICE, technologie | Classe lexicale 5 = Histoire comparée ou internationale | Classe lexicale 6 = Formation des enseignants |
|--|---------------------------------------|--|---|--|--|---|---|--|
| Échelle | | Macro | Micro/Macro | | Micro | Méso | Macro | |
| M3 Mots du sommaire du Traité dans Hal-SHS | 1182 Publi. | Emploi Marché | Infirmier Validation Vae Soignants | | Ergonomie Apprentissage Cognitif Activité (Ty_These) | Informatique Éducation Technologie Ordinateur | Siècle Histoire Social | Filtrée |
| M2 Explicite « FdA » Dans Hal-SHS | 172 Publi | ANI Réforme Chômage État | Sentir Lieu Acteur Société Demander | Acquérir Certification Validation Femme | Peu de travaux sur « apprentissages » et de « langues » dans le corpus identifié par M2 | Technologie Distance Formateur TICE | Peu de travaux sauf pour les docs qui ne sortent pas aussi dans M3 | Filtrée |
| M1bis Toutes les thèses du Sudoc | 275 Thèses | (<2000) Emploi Entreprise Cas Filière Employeur | Sensible Penser Sujet Épistémologique | | Apprenant Langue Résultat Adulte | Peu de thèses sur les « TICE » pour adultes | (<2000) Pays Éducation Priorité | Filtrée |
| M1 Thèses 2014-2015 du Sudoc | 41 Thèses | Entreprise Transfert Technologie | Sujet Logique Identitaire Professionnel Dynamique | | Apprentissage Apprenant Linguistique Adulte | Peu de thèses récentes sur les « TICE » pour adultes | Peu de thèses récentes de comparaison internationale ou historique | Enseignant, Form. continue Classe Universitaire |

Tableau T2. Comparaison des classes de mots repérées par le logiciel Iramuteq dans les publications de recherche concernant la formation des adultes selon la requête et la base de données

3. Examen complémentaire des points communs et des spécificités des différents corpus

L'examen de ce tableau T2 suggère d'effectuer une nouvelle analyse qui inclurait à la fois les trois dans une cartographie factorielle commune. Elle permettrait de confirmer nos résultats et de les visualiser de manière plus directe.

3.1. Similitude des grandes classes lexicales générées

Pour ce faire ont été regroupés dans un seul fichier tous les résumés des documents repérés dans les trois corpus M1bis, M2 et M3 (sans pour autant enlever les doublons pouvant exister entre M1bis et M2 ou M1bis et M3). La figure 1a donne le résultat d'une classification lexicale sur ce nouvel ensemble, comparé à celui de M3 seul, reproduit en figure 1b. Ces deux arborescences ont des architectures globalement similaires en trois noyaux centraux (à l'ordre des classes sur la figure près, ce qui n'est pas significatif avec Iramuteq). Pour l'une comme pour l'autre, on trouve les deux classes macro-sociales (« emploi, marché, économie, politique, entreprise » d'un côté et « histoire sociale » de l'autre), la classe « TICE et pédagogie » et enfin deux classes imbriquées regroupant « certification et validation » et « didactique » associée aux professions de santé².

² Il est tout de même à noter que si nous n'avions pas filtré très fortement les travaux sur la formation des maîtres (enseignement primaire, secondaire et supérieur), nous aurions obtenu au moins une classe supplémentaire dédiée à ce sujet.

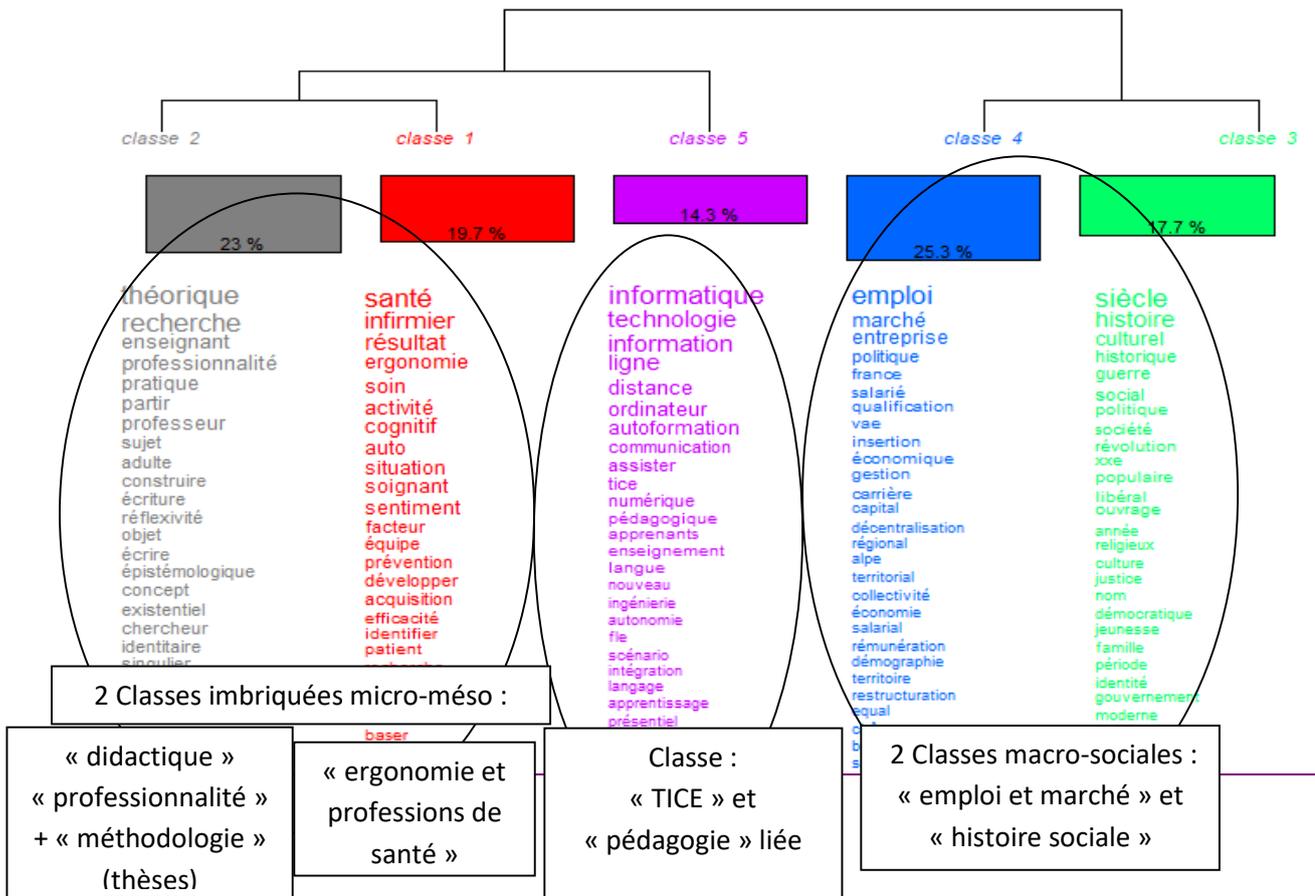


Figure 1a. Dendrogramme du corpus global obtenu par le cumul des 3 méthodes (M1bis + M2 + M3)

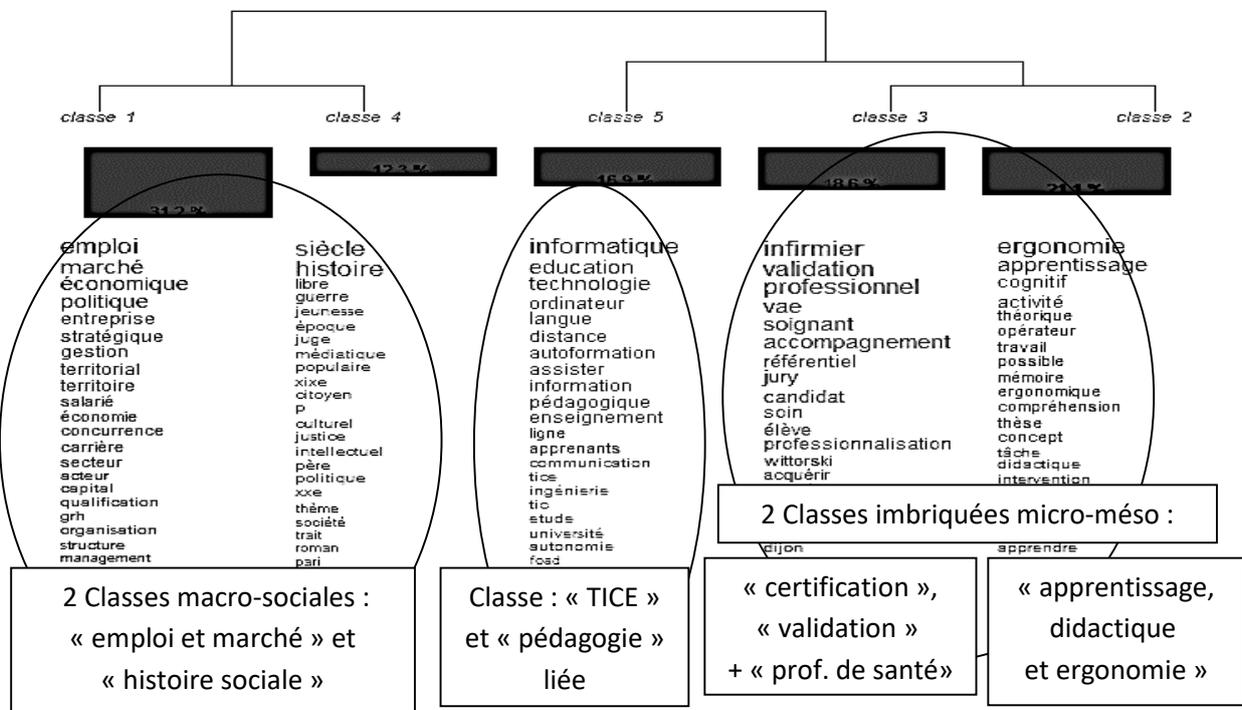


Figure 1b. Rappel du dendrogramme du corpus obtenu par la seule méthode M3

du corpus issu de M3 seul (voir les graphiques de l'article n° 3 que nous ne reproduisons pas ici faute de place).

3.3. Choix de la méthode (M1bis +M2+M3) comme référence

Ces résultats nous conduisent à considérer aujourd'hui cette méthode de recherche cumulant M1bis+M2+M3 comme un outil de référence pour les analyses ultérieures des publications de recherche en formation des adultes.

Dans la suite, l'idée sera de répéter ce cumul des 3 méthodes, par exemple année après année puis d'effectuer une analyse lexicale du corpus obtenu. Les résultats seront comparés pour voir comment vont progressivement se déformer ces cinq classes et ce triple noyau central (perspective macro, perspective micro/méso et TICE et pédagogie lié) ; ces déformations seront analysées ensuite comme des indicateurs des tendances à l'œuvre dans le déploiement de la recherche en FdA.

4. Alimentation des classes lexicales par les méthodes et sous-requêtes

D'ores et déjà, l'analyse Iramuteq nous permet aussi d'identifier parmi les corpus issus de M1bis, M2 et M3 (et même parmi les sous-requêtes de M3) ceux qui contribuent le plus à alimenter plus spécifiquement telle ou telle classe lexicale, ce qui revient à affiner les observations du tableau T2. Comme le montre la figure 3, on repère alors les quatre spécificités suivantes :

- La classe « didactique », « professionnalité » et le vocabulaire méthodologique qui s'y est agrégé sont surtout alimentés par les thèses du Sudoc (requête M1bis).
- La classe des professions de santé est surtout alimentée par la méthode M3 et plus précisément par les sous-requêtes correspondant aux chapitres « didactique pro », « apprentissage en situation de travail » et « ergonomie » du Traité.
- La classe « TICE » et « pédagogie » liée est aussi particulièrement alimentée par la méthode M3, mais spécifiquement par les sous-requêtes correspondant aux chapitres « autoformation » et « ingénierie pédagogique » du Traité.
- La classe « histoire sociale » se révèle plus spécifique aux documents repérés par la méthode M2 comme mentionnant la « formation des adultes » dans Hal-SHS mais ne correspondant à aucun chapitre du Traité, sans doute parce que les sujets traités sont trop éloignés des questions vives actuelles.

Enfin, on observe que la cinquième classe (« emploi », « entreprise ») est alimentée conjointement par toutes les méthodes sans spécificité particulière.

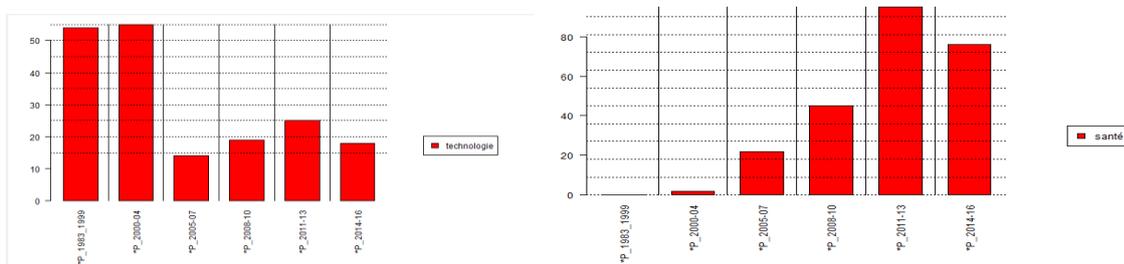
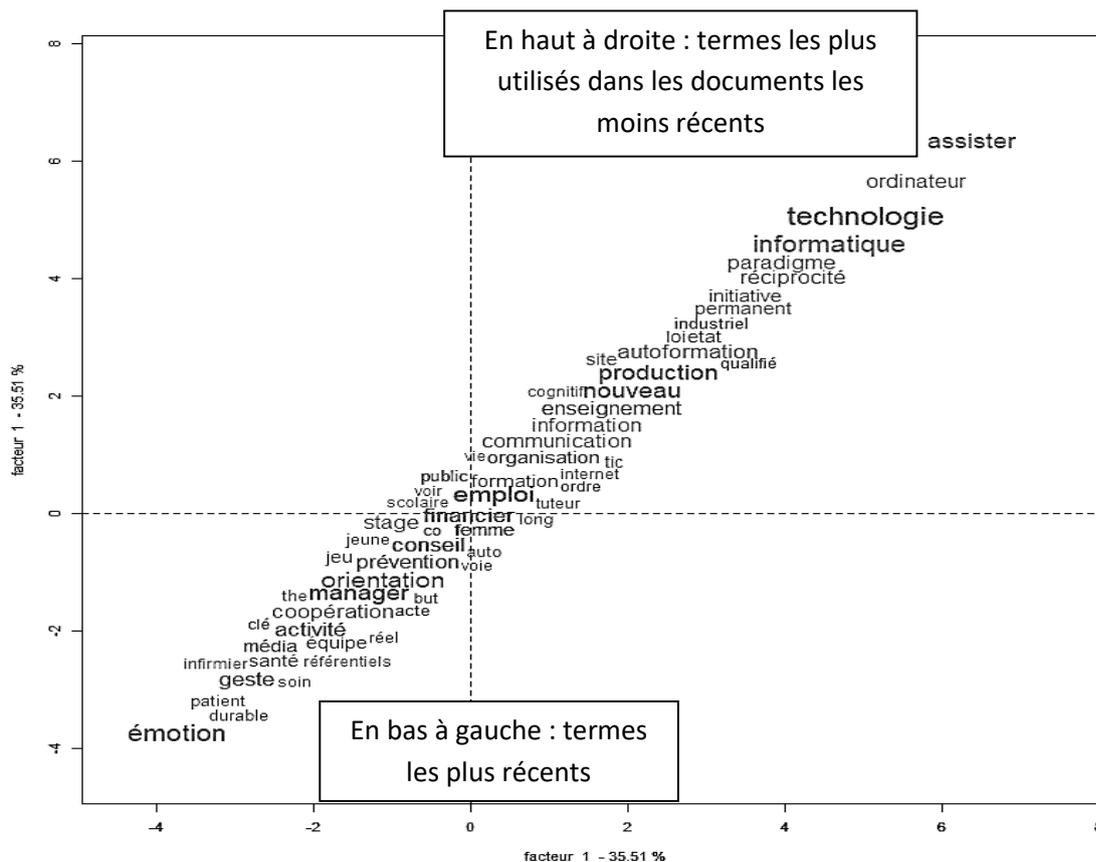
5. Conclusion

Cet article peut être lu à deux niveaux : d'une part, il fournit plusieurs visions de la production actuelle des publications de recherche et de leurs spécificités ; d'autre part, il répond à un problème méthodologique au cœur de cette rubrique : le fait que dans Hal-SHS comme dans le Sudoc la plupart des publications qualifiables de « recherche en formation des adultes » sont indexées bien trop finement pour pouvoir être retrouvées par une seule requête ou méthode générale.

Sur ce point, il avait donc semblé nécessaire de multiplier diverses méthodes et séries de sous-requêtes spécifiques pour ensuite se faire une vision par triangulation de l'évolution du champ. Ce que cet article permet de constater, c'est que l'addition simple des trois méthodes de recherches que sont M1bis, M2 et M3 avec toutes les sous-requêtes, suivie d'une classification Iramuteq du corpus global, permet d'aboutir à une architecture en classes lexicales lisible, tout en conservant une traçabilité des diverses

Encadré : évolution temporelle des termes fréquents

Le corpus global peut être utilisé pour donner une vision d'ensemble des variations temporelles des termes utilisés. Les résultats en sont présentés dans les figures 4, 5a et 5b. Concrètement, la figure présente en diagonale l'axe 1 d'une analyse factorielle des mots utilisés plus de 35 fois dans les résumés du corpus. La période a été découpée en six tranches : avant 1999, 2000-2004, 2005-2007, 2008-2010, 2011-2013 et 2014-2016. Plus un terme apparaît en bas à gauche, plus sa fréquence s'est développée dans le corpus. Plus il est en haut à gauche, plus il s'est raréfié. En 5a et 5b, sont présentés à titre d'exemple les nombres d'occurrences d'un mot plus récemment utilisé (santé) et inversement un moins contemporain (technologie). On observe que le vocabulaire et sans doute les thématiques sont devenus plus focalisés. On ne parle plus de « ordinateur » ou de « technologie » en général et on s'intéresse aujourd'hui aux professions de santé, à l'émotion, aux activités, aux gestes...



Figures 4, 5a et 5b. Comparaison temporelle des usages des termes du corpus global (axe 1 d'une analyse factorielle sur 6 périodes : antérieur à 1999, 2000-2004, 2005-2007, 2008-2010, 2011-2013 et 2014-2016). Plus un terme apparaît en bas à gauche, plus sa fréquence s'est développée dans les plus récentes années dans le corpus. Plus il est en haut à gauche, plus il s'est inversement raréfié. En 5a et 5b sont présentés à titre d'exemple les nombres d'occurrences d'un mot plus récemment utilisé (santé) et inversement d'un autre, moins contemporain (technologie) ; 2016 est à droite sur ces graphiques.

Références bibliographiques

Carré P. et Caspar P. (2011). *Traité des sciences et techniques de la formation* (3^e éd.). Paris : Dunod.

Las Vergnas O. (2016, 2017) [n° 1, n° 2 et n° 3]. Rubrique « Vie de la recherche ». *Savoirs*, n° 40, 41 et 42, Paris, L'Harmattan ; et revue *TransFormations*, Université de Lille.

Ratinaud P. et Déjean S. (2009). IRaMuTeQ : implémentation de la méthode ALCESTE d'analyse de texte dans un logiciel libre. Communication au congrès « Modélisation Appliquée aux Sciences Humaines et Sociales » (MASHS2009). Université de Toulouse Le Mirail.