



HAL
open science

Analyse diachronique du processus de déterminologisation. Une réflexion en diachronie courte en physique des particules

Aurélie Picton, Anne Condamines, Julie Humbert-Droz

► **To cite this version:**

Aurélie Picton, Anne Condamines, Julie Humbert-Droz. Analyse diachronique du processus de déterminologisation. Une réflexion en diachronie courte en physique des particules. Cahiers de Lexicologie, 2021, 1 (118), pp.193-225. 10.48611/isbn.978-2-406-12006-3.p.0193 . halshs-03344283

HAL Id: halshs-03344283

<https://shs.hal.science/halshs-03344283>

Submitted on 14 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**ANALYSE DIACHRONIQUE DU PROCESSUS DE
DETERMINOLOGISATION**

Une réflexion en diachronie courte en physique des particules

Cet article propose une réflexion sur la diachronie en langues de spécialité, à travers la question du processus de déterminologisation. Sur la base d'une analyse d'un corpus diachronique complexe, comparable en genres et degrés de spécialité, nous discutons les fonctionnements en jeu dans ce processus, en diachronie courte. Notre analyse permet de mettre en lumière les rôles des sous-corpus dans ce processus, ainsi que la spécificité des dynamiques dans le domaine de la physique des particules.

This paper is a reflection on diachrony in languages for special purposes, through the process of determinologisation. On the basis of an analysis of a complex diachronic corpus, comparable in genres and levels of specialisation, we discuss the linguistic phenomena involved in this process, in short diachrony. Our analysis sheds light on the roles of sub-corpora in this process, as well as on the specificity of these dynamics in the field of particle physics.

Mots-clés : diachronie courte, déterminologisation, corpus spécialisé, corpus comparable

Introduction

Dans cet article, nous souhaitons alimenter la réflexion sur la diachronie en langues de spécialité, à travers l'analyse, outillée en corpus, d'un phénomène spécifique : la déterminologisation. Plusieurs auteurs ont montré l'importance de la perspective diachronique dans l'analyse de ce phénomène (p. ex. Dury 2008, Renouf 2017). Dans ces travaux, les données reposent généralement sur une comparaison de corpus spécialisé et général (de presse), pour mettre en évidence l'évolution des fonctionnements des termes dans ces deux ensembles. Cependant, la déterminologisation en tant que processus de passage des termes dans différents genres et degrés de spécialité et dans le temps (Humbert-Droz *et al.* 2019) requiert de

chercher à décrire plus finement le fonctionnement des unités dans ce double processus.

Nous proposons donc ici un regard en diachronie courte (entre 2003 et 2016) sur la déterminologisation en physique des particules, dans plusieurs genres et degrés de spécialité. Différentes questions se posent, qui touchent à la dynamique des fonctionnements des unités dans les différents sous-corpus. En particulier, ce type d'analyse permet de repérer la manière dont les fonctionnements des unités qui se déterminologisent évoluent dans chacun des sous-corpus à l'étude. L'objectif est ainsi de mettre en évidence des « traces » linguistiques du passage de ces unités, et potentiellement des parallélismes, des évolutions temporelles communes (voire successives) entre les différents genres/degrés de spécialité. Dans ce contexte, la prise en compte de la dimension temporelle offre un portrait plus complet des dynamiques en jeu dans les relations entre déterminologisation et diachronie, et ouvre plusieurs questionnements théoriques et méthodologiques sur la diachronie.

Nous proposons de discuter ces différents aspects en cinq temps. Tout d'abord, nous définissons notre cadre théorique de travail (section 1). Nous présentons ensuite notre cadre méthodologique, qui repose sur un corpus comparable à la fois en diachronie courte et entre genres et degrés de spécialité. La question du découpage temporel de notre corpus diachronique est une question essentielle pour assurer la comparaison multiple des sous-corpus, tout en respectant la dynamique des données. Nous discutons ce point dans la section 3. Nous développons nos analyses dans la section 4, avec un accent particulier sur les variantes dénominatives et la distribution de termes désignant les particules. Nous terminons par une discussion et quelques remarques conclusives.

1. Contexte théorique

1.1. Diachronie et terminologie

En langue générale, les approches diachroniques apparaissent très variées et soutenues par les progrès et apports de la linguistique de corpus (Gries et Hilpert 2008, Prévost 2015). Cette vitalité permet aujourd'hui d'avancer et de mûrir des

réflexions théoriques et méthodologiques, dont les langues de spécialité peuvent tirer profit.

En effet, en langue de spécialité, la diachronie a longtemps été décrite comme « le parent pauvre » des recherches scientifiques (pour différentes raisons historiques et théoriques discutées entre autres par Dury et Picton 2008, Humbley 2011, Møller 1998, Van der Yeught 2009). Avec le changement de paradigme amorcé dans les années 1990 par les théories de la terminologie, la diachronie a repris sa légitimité, mais a gardé jusqu'à récemment (Dury 2013) une place assez confidentielle. Aujourd'hui, la notion de « diachronie » fait l'objet de réflexions et discussions importantes quant à son rôle et son statut en terminologie (p. ex. Humbley, 2018 : 203 *sqq.*, Dury 2018, Picton 2018). Alors qu'elle est présentée parfois comme un type de variation dialectale (p. ex. chez Freixa 2006 ou Faber et León-Araúz 2016), la terminologie diachronique est décrite comme un « champ d'investigation » à part entière par d'autres auteurs (Dury 2013 : 2, 2018). Cette deuxième perspective permet d'ouvrir les réflexions sur la variation dans les langues de spécialité, et la circulation des termes, dans le temps. Dans cet article, c'est donc la diachronie comme perspective pour l'analyse des corpus spécialisés qui est discutée, et en particulier la diachronie dite « courte » (Siouffi *et al.* 2012, Picton 2014, Banks 2016).

1.2. Circulation des termes : le cas de la déterminologisation

La déterminologisation fait appel à la notion de circulation des termes, qui trouve ses origines dans les travaux de Guilbert. À travers la question de la néologie, l'auteur constate que la création de termes peut passer par des transferts de termes entre langues de spécialité ou entre langue générale et langues de spécialité. De même, certains termes ont tendance à « passer » dans la langue générale (Guilbert 1965, 1973, 1975 notamment). Ces travaux ont suscité d'importantes réflexions à propos des mouvements permanents entre langues de spécialité et langue générale (p. ex. Gaudin 2003, Delavigne 2020, Calberg-Challot 2007, Ungureau 2006, Dury 2008). La notion de circulation

des termes s'est toutefois affirmée à travers le prisme de la socioterminologie et a été largement explorée par des auteurs tels que Guespin (1991), Gambier (1991), Gaudin (1993, 2003) ou encore Delavigne (2001). Ces auteurs se sont ainsi consacrés à la description des fonctionnements des termes dans la diversité des situations de communication dans lesquelles ils sont utilisés lorsqu'ils circulent en dehors des discours des spécialistes, et en particulier dans le cadre de la vulgarisation (Delavigne 2001, 2003, 2020).

Parallèlement, d'autres travaux portant sur la circulation des termes ont abordé cette question du point de vue de la déterminologisation (p. ex. Ungureanu 2006, Dury 2008 ; Renouf 2017), faisant ainsi suite aux réflexions de Meyer et Mackintosh (2000). C'est dans cette perspective que nous nous situons.

Le terme *déterminologisation*, proposé par Meyer (Meyer, Mackintosh et Varantola 1998, Meyer 2000, Meyer et Mackintosh 2000), renvoie aux mouvements de termes qui se produisent d'une langue de spécialité vers la langue générale. Dans ces travaux, un point de vue essentiellement sémantique est adopté, les autrices s'intéressant particulièrement à « l'étirement sémantique qui se produit lorsqu'un terme attire l'attention du public » (Meyer et Mackintosh 2000 : 199). Les termes sont alors appréhendés à partir de leur fonctionnement dans la langue générale. Or la déterminologisation peut également être considérée comme un processus d'intégration de termes dans la langue générale, qui fait intervenir deux dimensions.

Pour ce qui est de la première dimension, il semble évident que les termes ne passent pas d'une langue de spécialité à la langue générale directement, principalement pour deux raisons. D'une part, comme l'ont montré les travaux en socioterminologie traitant de circulation des termes, dès lors que les termes sortent de la sphère spécialisée, ils peuvent s'inscrire dans des discours tout à fait variés, sans pour autant que ces discours relèvent de la langue générale (l'exemple le plus parlant est la vulgarisation scientifique). D'autre part, si nous admettons que les langues de spécialité et la langue générale ne forment pas deux ensembles clos mais un continuum, à l'instar de Cabré (1998 : 124) ou

Gaudin (1993 : 130), alors nous pouvons considérer que la variété des discours dans lesquels les termes sont utilisés peuvent se situer à divers endroits dans ce continuum.

Pour ce qui est de la seconde dimension, et comme l'ont montré notamment Dury (2008) et Renouf (2017), la déterminologisation est un processus qui se produit nécessairement dans le temps. Il est ainsi admis que non seulement les termes intègrent progressivement la langue générale, mais également qu'ils voient leur sens se modifier au cours du temps, au gré de leurs usages dans divers types de textes et à des fins variées.

Pour ces raisons, nous considérons que la déterminologisation renvoie à un processus de passage de termes d'une langue de spécialité vers la langue générale, qui s'effectue progressivement, à la fois dans le temps et dans le continuum entre langues de spécialité et langue générale.

L'analyse de la déterminologisation dans un domaine spécialisé nécessite une réflexion particulière sur la constitution du corpus d'étude. Il faut en effet supposer une circulation des termes dans des situations qui vont de la plus spécialisée à la moins spécialisée, cette dernière pouvant servir, par hypothèse, de lieu d'arrivée de ces termes dans la langue générale. On peut donc considérer que le processus fait intervenir essentiellement deux communautés discursives (Swales, 1990) : les experts du domaine et les médiateurs (journalistes le plus souvent) qui sont en relation avec des locuteurs non experts du domaine. Entre les deux communautés, un ensemble de situations contribue à la diffusion des termes vers un usage de moins en moins spécialisé. Nous avons fait l'hypothèse d'une chaîne de dialogues entre différents groupes de locuteurs, plus ou moins liés aux experts ou/et aux médiateurs, dont les discours s'inscrivent dans différents genres textuels (voir 2.2).

2. Contexte méthodologique¹

2.1. Domaine

¹ La réflexion proposée ici repose sur le travail et les données constituées par Julie Humbert-Droz dans le cadre de sa thèse (en cours).

Le domaine auquel nous nous intéressons est celui de la physique des particules. Ce choix répond à certaines caractéristiques, propres à la terminologisation. En effet, il est reconnu que les termes qui s'intègrent de manière privilégiée dans la langue générale sont des termes qui appartiennent à des domaines qui sont soit largement médiatisés, ce qui inclut aussi bien les « médias ordinaires » (Moirand 1999 : 5) que les médias de vulgarisation, soit proches du quotidien des locuteurs (Guilbert 1975 : 84, Meyer 2000 : 46, Gaudin 1993 : 131, Moirand 2007 : 20 parmi d'autres). La physique des particules correspond à cette situation, non seulement car c'est un domaine qui apparaît régulièrement dans les médias (ordinaires et de vulgarisation), mais également car l'Organisation européenne pour la recherche nucléaire (CERN), organisme de recherche particulièrement important dans ce domaine², se situe à la frontière franco-suisse. Ceci implique un important volume de publications, dont certaines sont dédiées à la population environnante, et que nous pouvons exploiter dans une démarche en corpus.

2.2. Corpus et méthode

L'analyse mise en place s'inscrit dans la perspective de la terminologie textuelle. Sans rentrer dans les détails ici, (voir Condamines et Picton, à paraître), nous pouvons retenir quatre points caractérisant cette approche :

- l'analyse est réalisée sur un corpus dont la constitution est minutieusement réfléchi en fonction de l'objectif de l'étude,
- les termes sont le point de départ de l'analyse, même lorsque l'objectif de l'étude concerne plutôt l'aspect discursif ; l'exploration du contexte distributionnel des termes est alors fondamentale,
- l'analyse s'appuie sur des résultats d'outils d'exploration des données sélectionnés spécifiquement pour l'analyse,

² Le LHC se situe au CERN et le boson de Higgs y a été découvert (voir 2.2).

- des experts du domaine sont sollicités tout au long de l'étude (constitution du corpus, validation des résultats, etc.).

Le corpus constitué pour notre étude est un corpus comparable composé de cinq sous-corpus (SC), pensés pour refléter la diversité des genres textuels et des degrés de spécialité impliqués dans le processus de déterminologisation (voir 1.2 et Humbert-Droz *et al.* 2019). En fonction des supports de diffusion de l'information disponibles, ces SC sont composés comme suit :

- un SC *Spécialisé* (SC-SPE), incluant des articles spécialisés et des thèses de doctorat, composé de 994 875 occurrences ;
- un SC *Communiqués* (SC-COM), composé de communiqués de presse et comptant 210 320 occurrences ;
- un SC *Rapports* (SC-RAP), qui contient les rapports d'activités annuels de différents laboratoires de recherche. Il se compose de 1 141 873 occurrences ;
- un SC *Vulgarisation* (SC-VULG), composé d'articles de revues et de sites Web de vulgarisation. Il compte 620 045 occurrences ;
- un SC *Presse* (SC-PRESSE), incluant des articles de presse généraliste, pour représenter non pas la langue générale, mais un genre non spécialisé participant du processus de déterminologisation. Il compte 1 098 708 occurrences.

Le corpus couvre la période de 2003 à 2016. Le choix de cette période résulte d'une réflexion, menée conjointement avec des expertes du domaine, dont le but consistait à identifier des événements-clés de la physique des particules. Nous avons en effet fait l'hypothèse que des événements importants de ce domaine (découvertes, avancées technologiques remarquables, etc.) pourraient influencer la manière dont les termes en lien

avec ces événements sont utilisés dans divers textes. Ces influences seraient alors reflétées dans des changements linguistiques observables en corpus.

À l'issue de cette réflexion, deux événements principaux ont été identifiés, qui sont la mise en marche du LHC (*Large Hadron Collider*) en 2008 et l'annonce officielle de la découverte expérimentale du boson de Higgs en 2012. Le choix de la période 2003-2016 est alors contraint par la nécessité de recueillir suffisamment de données avant/après ces deux événements, rendant les changements linguistiques observables dans le corpus. Pour ce faire, des textes ont été compilés pour chaque année de cette période, pour chaque SC.

Nos analyses sont basées sur un échantillon de termes obtenu à partir du corpus et avec l'aide des expertes. Plus précisément, l'extraction des candidats-termes à partir du corpus a montré que le terme *particule* était particulièrement fréquent dans tous les SC, ce qui s'explique par le fait que les particules sont les objets d'étude du domaine. Puisque les termes qui se déterminologisent sont ceux qui « attire[nt] l'attention du public » (Meyer et Mackintosh, 2000 : 199), nous pensons que les termes désignant les particules sont susceptibles d'apparaître dans les médias et donc d'intégrer la langue générale. Nous avons décidé de focaliser l'analyse sur les particules, à partir de listes fournies par les expertes et complétées avec les données du corpus (219 termes au total). Par ailleurs, selon les expertes, les particules peuvent être classées en deux types : les particules prouvées, qui ont été observées expérimentalement, et les particules hypothétiques, qui sont uniquement prévues par la théorie.

Pour l'observation, nous avons travaillé avec des concordanciers (AntConc (Anthony 2018) et Sketch Engine (Kilgarriff *et al.* 2014)), un extracteur de termes (TermoStat (Drouin 2003)), un analyseur syntaxique (Talismane (Urieli 2013)) et des scripts Perl et R (R Core Team 2018) ponctuels.

3. Découpage diachronique des données

En diachronie, le découpage temporel des données est une question cruciale, mais parfois difficile à trancher. La littérature rend compte de deux stratégies :

- soit le corpus constitué est organisé en SC en fonction d'un événement extralinguistique marquant, identifié *a priori* (p. ex. Drouin *et al.* 2006),
- soit le corpus constitué est « découpé » temporellement sur une base régulière, en fonction du « grain » d'analyse souhaité ; Dury (2008) par exemple, dans une analyse sur 10 ans, propose un découpage des SC tous les deux ans, afin de construire une analyse très fine ; Ledouble (2019) quant à elle travaille sur un découpage annualisé.

3.1. Critères de découpage

Ces deux stratégies ont des atouts, mais elles présentent plusieurs limites, que notre contexte d'étude éclaire. Dans notre corpus, nous pourrions en effet découper nos données sur la base d'un avant/après des événements identifiés comme marquants par nos expertes lors du choix de la période 2003-2016. Néanmoins, ce choix a un impact sur nos hypothèses et notre analyse, puisqu'il implique de postuler l'existence d'un lien entre la date de ces événements et la déterminologisation des termes du domaine, dans tous les SC. Ceci est possible, mais impose un biais que nous ne souhaitons pas apporter *a priori*. En particulier, puisque l'objectif de notre recherche est de saisir les liens entre diachronie et déterminologisation, l'usage de certains termes auprès de certains groupes de locuteurs pourrait ne pas concorder avec le découpage prédéfini.

À l'instar d'autres chercheurs, nous pourrions également découper notre corpus sur la base d'intervalles temporels arbitraires. Néanmoins, plusieurs limites se posent :

- d'une part, nous ne disposons pas de critères objectifs pour décider de l'empan temporel à choisir. S'il est clair que plus un empan temporel est court, plus l'analyse sera fine (Dury 2008), un intervalle de 2 ans est-il vraiment différent d'un intervalle d'1 an ?
- D'autre part, un découpage régulier mais arbitraire peut générer du silence dans les résultats obtenus, si une évolution ne correspond pas au découpage choisi.

- Enfin, puisque nous prenons en compte, dans l'analyse, différents genres et degrés de spécialité associés à différents groupes de locuteurs, nous devrions découper chacun des cinq SC avec le même empan. L'observation des SC ainsi constitués (35, si l'on découpe tous les 2 ans) devient extrêmement difficile. Il devient en effet quasiment impossible techniquement et cognitivement de comparer autant de données et faire émerger des comportements linguistiques fiables. Ce type de découpage n'est donc envisageable ici que pour des observations quantitatives (fréquences et répartition).

Pour ces différentes raisons, afin d'assister notre choix de découpage, nous explorons une proposition méthodologique élaborée en linguistique diachronique : l'utilisation de méthodes statistiques pour faire émerger des corpus eux-mêmes des regroupements diachroniques. L'approche choisie ici est développée par Gries et Hilpert (2008, 2012). À travers une classification ascendante hiérarchique par contiguité (*variability-based neighbor clustering, VNC*), les auteurs proposent de prendre en compte certaines caractéristiques des unités étudiées (p. ex. leur fréquence, leur score d'association, etc.) dans des corpus découpés en tranches successives pour identifier des regroupements de documents possédant des propriétés communes. Ils dégagent ainsi automatiquement de ces données, de manière endogène, un découpage, qui se base sur une certaine homogénéité de fonctionnement³. Nous proposons ici de mettre en regard ce découpage avec nos connaissances (exogènes) des événements du domaine.

3.2. Découpage assisté par la statistique : observations

Nous l'avons dit, dans cet article, nous proposons de focaliser notre réflexion sur un groupe de termes spécifiques, désignant

³ Et qui reposent ici sur le coefficient de corrélation de Pearson (« Pearson product-moment correlation coefficient »).

les particules (voir 2.2). L'approche VNC est donc appliquée à cette liste de 219 termes et leur fréquence dans chaque SC. Les résultats obtenus se présentent sous forme de regroupements (clusters) hiérarchiques (Figures 1 et 2).

Figure 1 – VNC « particules », SC-SPE (gauche) et PRESSE (droite)

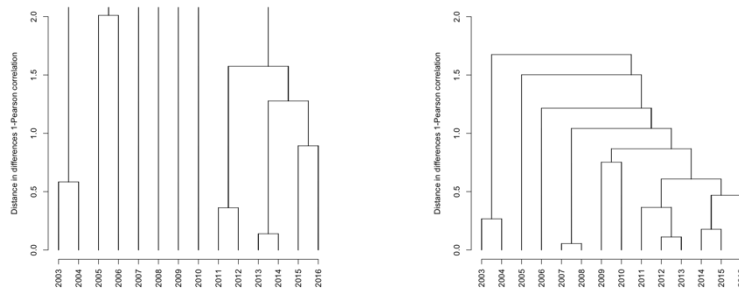
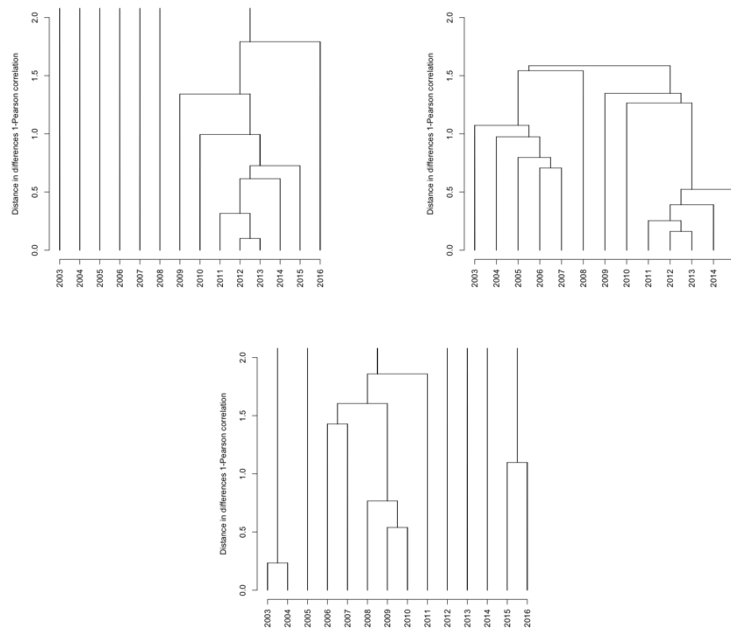


Figure 2 – VNC « particules », SC-COM (gauche), RAP (droite) et VULG (dessous)



Nous nous concentrons dans un premier temps sur les tendances globales qui se dégagent de ces résultats.

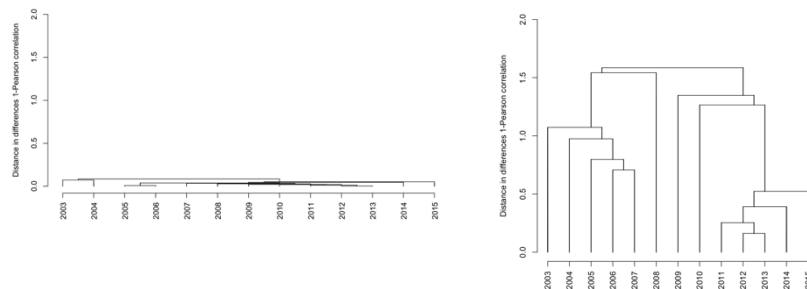
On observe que, pour chacun des SC testés, des regroupements assez différents apparaissent. Ainsi, la Figure 2 montre globalement autant de dynamiques différentes en diachronie que de SC :

- les rapports (SC-RAP) semblent montrer une dynamique globale en deux périodes assez homogènes 2003-2007 et 2011-2016,
- les communiqués (SC-COM) en une période homogène (2009-2016)
- et les textes de vulgarisation (SC-VULG) également en une période homogène, mais distincte des communiqués (2006-2011).

Cette approche par VNC montre donc des dynamiques différentes sur un même groupe de termes, dans ce domaine.

Pour aller plus loin, la Figure 3 met en regard les résultats obtenus dans le SC-RAP, sur les unités désignant les particules et les 33 unités désignant des types d'accélérateurs de particules.

Figure 3 – VNC « accélérateurs de particules » (gauche) vs « particules » (droite) (SC-RAP)



Les deux classifications obtenues indiquent que ces deux groupes de termes ont des comportements différents entre 2003 et 2016. La dynamique des termes désignant les particules, dans ce SC, se dessine en deux périodes homogènes (à droite), alors qu'une analyse des termes désignant les accélérateurs ne fait

plus apparaît ces deux périodes (à gauche). Cette impression peut être confirmée par la nature même de ce SC : en effet, la structure de chaque rapport comporte une section spécifique aux actualités des accélérateurs de particules, alors qu'aucune section relative spécifiquement aux particules n'est prévue. Ainsi, la classification reflète la structure des documents compilés et la dynamique qui leur est propre.

Ces regroupements, assez généraux, suggèrent effectivement que l'hypothèse de dynamiques différentes en fonction des SC, et des termes observés, mérite d'être considérée pour les analyses diachroniques (section 3.1). Nous retenons donc à ce stade l'importance de ce facteur dans la prise de décision du découpage temporel des données diachroniques.

Pour pouvoir comparer nos cinq SC, nous cherchons à identifier des périodes récurrentes, où des regroupements statistiquement significatifs sont relevés. Nous proposons ici de nous focaliser sur les regroupements (clusters) les plus significatifs (Figure 1), c'est-à-dire ceux dont le coefficient de Pearson est le plus proche de 0 (et inférieur à 0.5), indiqué en ordonnée.

Dans nos données, une période semble ressortir de manière régulière, la période 2011-2013. Ainsi,

- dans le SC-PRESSE, un cluster significatif peut être repéré en 2011-2013,
- dans le SC-SPE, un cluster 2011-2012 et 2013-2014,
- dans le SC-COM et le SC-RAP, un cluster 2011-2013⁴.

L'intérêt ici est de repérer un équilibre entre les indications fournies par les regroupements endogènes et des événements marquants dans l'histoire de la physique des particules (indicateurs exogènes, voir 2.2), pour adapter notre découpage. Du point de vue de l'histoire du domaine, 2011-2013 est bien une période charnière, puisqu'elle correspond à la découverte du

⁴ Dans le cas du SC-VULG, il s'agit plus d'une « rupture » de clusters à partir de 2010, bien que nous ne les prenions pas directement en compte, puisque non significatifs en fonction des critères statistiques que nous avons établis.

boson de Higgs, annoncée le 4 juillet 2012 et confirmée le 15 mars 2013. Néanmoins, les clusters permettent d'affiner l'analyse de cette dynamique en mettant en évidence l'année 2011, qui, si elle ne correspond pas à une annonce officielle, correspond à une année essentielle pour les spécialistes qui distillaient progressivement l'information de cette découverte (qui ne s'est évidemment pas produite en un jour).

Ces réflexions sur le découpage endogène et exogène des données nous amènent à proposer deux découpages, complémentaires, pour l'analyse :

- un découpage annualisé de 2003 à 2016, dédié à l'observation de courbes de fréquences (cf. 3.1.),
- un découpage en deux périodes : 2003-2010 et 2011-2016.

4. Observations

4.1. Analyse globale quantitative

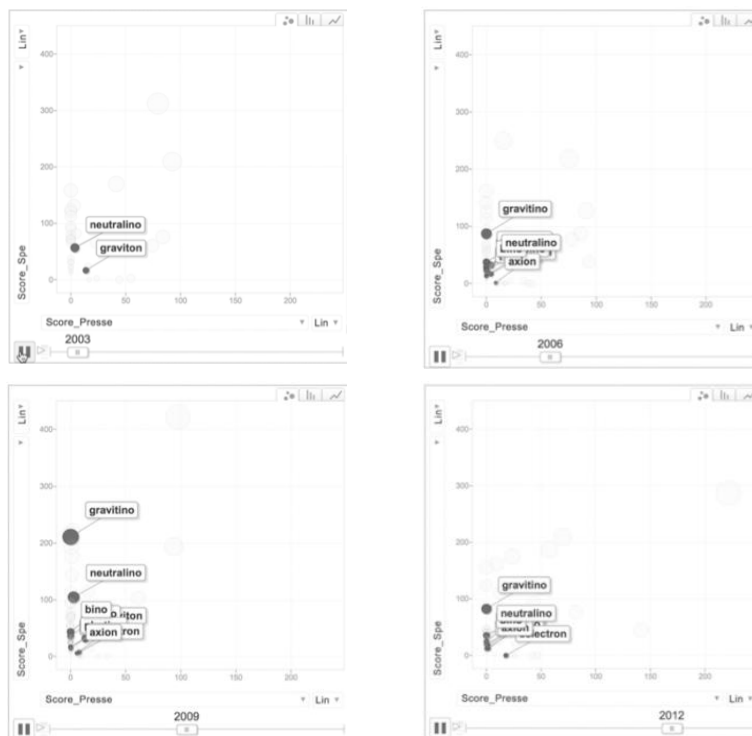
L'analyse des fréquences des termes dans les SC est un point d'entrée qui permet de mettre en évidence des tendances d'usages, entre les différents SC et dans le temps.

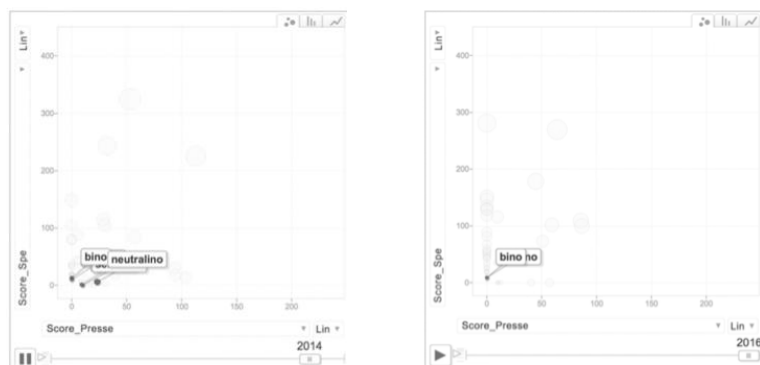
Nous adoptons ici deux stratégies : le test du Chi² et la visualisation en *Motion Charts* (Gesmann et De Castillo 2011).

Le score de Chi² permet d'observer les variations de fréquence statistiquement significatives entre les SC et en diachronie. Appliquée à notre liste de termes, la comparaison des périodes 2003-2010 vs 2011-2016 met en évidence différentes tendances. D'une part, le seul terme significativement attesté dans 4 des 5 SC en 2011-2016 est *boson de Higgs*. Sa fréquence n'est cependant pas significative dans le SC-SPE. Cette observation peut être mise en lien avec nos propos *supra* sur les dynamiques propres à chaque SC. Ici, une hypothèse serait que pour les spécialistes, le travail sur le boson de Higgs était en cours depuis de nombreuses années et sa découverte pressentie bien avant son annonce officielle par le CERN. Le score de Chi² non significatif pour *boson de Higgs* dans la seconde période couverte par ce SC indique donc que la découverte de ce boson n'est pas un événement particulièrement marquant pour les spécialistes.

L'observation de la répartition temporelle des autres types de particules dans les SC montre une seconde tendance : certains SC contiennent moins de termes relatifs aux particules que d'autres. Typiquement, le SC-PRESSE est celui qui en contient le moins. Enfin, les particules hypothétiques montrent un fonctionnement différent de celui des particules prouvées : les graphiques dynamiques présentés en Figure 3 (Picton *et al.* à paraître) montrent la constance des particules hypothétiques (en gris foncé, étiquetés), qui « restent » majoritairement dans le SC-SPE (à gauche) et ne « passent » que de manière très anecdotique dans le SC-PRESSE (en bas). En gris clair, plusieurs particules (dont *boson de Higgs* et ses variantes) voient leur fréquence augmenter dans les deux SC, en particulier après 2009 et 2012.

Figure 3 - Motion Chart – Dynamique des particules hypothétiques vs prouvées – 2003-2016





Ces éléments viennent confirmer l'existence de fonctionnements (quantitatifs) différents, non seulement entre les SC, mais également entre les types de termes. De ce point de vue, il semble qu'une sélection s'opère lorsque les termes « passent » d'un SC spécialisé à un SC plus général, selon la nature des concepts en jeu. Ainsi, tous les types de particules n'attireraient pas l'attention du grand public de la même manière (voir 2.2), les particules prouvées semblant ici privilégiées.

4.2. Le cas de boson de Higgs

Pour aller plus loin, nous focalisons notre attention sur la seule unité statistiquement significative dans la majorité des SC : *boson de Higgs*, dont la particularité est que son existence a été prouvée en 2011-2013.

4.2.1 Fonctionnement des variantes dénominatives

En diachronie, l'analyse des variantes dénominatives est associée à de nombreux phénomènes linguistiques, tels que le foisonnement en situation de néologie (Guilbert 1973, 1975), la synonymie de circonstance (Dury et Lervad 2008), la centralité (Jacques 2005) ou encore l'impact d'une « concurrence dénominative » sur la compréhension des concepts (Ledouble 2019, 2020).

Dans notre corpus, 14 variantes dénominatives sont attestées : *boson de Higgs*, *boson de higgs*, *boson BEH*, *boson beh*, *particule Dieu*, *particule de Dieu*, *particule de dieu*, *particule de*

Higgs, boson de Brout-Englert-Higgs, boson scalaire de Higgs, boson H, boson h, Higgs et boson Higgs.

Table 1 – Répartition des 14 variantes dénomminatives

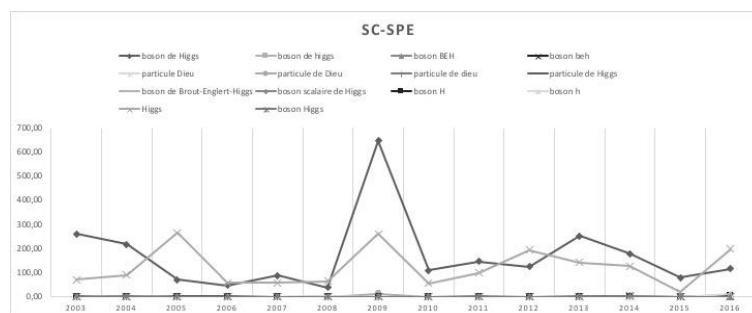
SC	SPE	COM	RAP	VULG	PRESSE
2003-2010	7	2	5	5	5
2011-2016	8	4	5	8	8
Variantes communes	6	2	3	4	4

Du point de vue quantitatif (Table 1), le SC-SPE est celui qui compte le plus de variantes possibles (6), mais la tendance reste stable dans le temps. Par contre, le nombre de variantes augmente significativement après 2011 dans les SC-COM, VULG et PRESSE, en lien très certainement avec la centralité soudaine de ce concept après 2011.

Du point de vue qualitatif (Figures 4 et 5), il ressort que les deux variantes préférées sont les mêmes pour tous les SC, sur toute la période considérée : *boson de Higgs* et *Higgs*.

Pour les autres variantes, plusieurs sont propres à certains SC et à certaines périodes. Ainsi *particule de Dieu* n'est attestée que dans le SC-PRESSE, *boson beh* dans SC-VULG ; *Higgs* baisse dans SC-VULG, qui préfère *boson H* après 2010, alors qu'elle devient la variante préférée dans SC-COM. Les autres variantes, partagées par 2 à 4 SC, montrent également une répartition propre.

Figure 4 – Évolution des variantes de *boson de Higgs* (2003-2016) - fréquences relatives



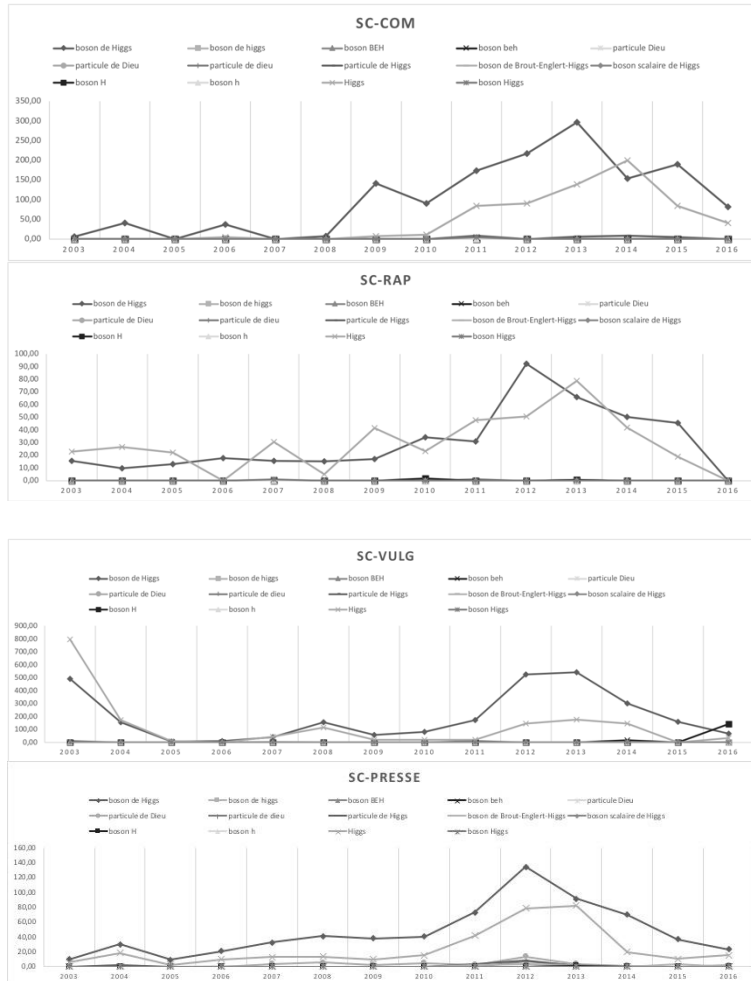
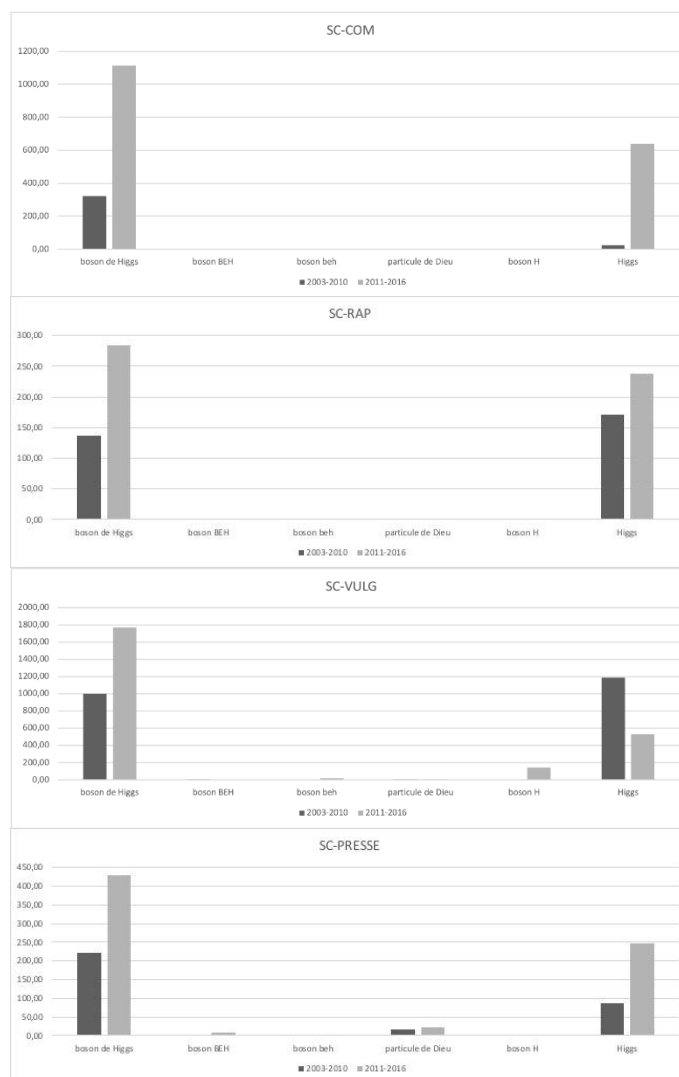


Figure 5 – Évolution des variantes (2003-2010 vs 2011-2016)





Chaque SC montre donc une sélection et une évolution de variantes qui lui sont propres. Dans le SC-SPE, le « pic » de fréquence de ces dénominations apparaît plus tôt (autour de 2009), alors que les autres SC montrent un pic sur la période 2011-2012. Cette observation rejoint les remarques formulées plus haut et amènent à s'interroger sur la succession des fonctionnements linguistiques. En effet, bien que certains phénomènes semblent apparaître plus tôt dans le SC-SPE (p. ex.

l'augmentation des occurrences de *boson de Higgs* et de ses variantes), il nous semble difficile d'affirmer que le SC-SPE aurait nécessairement influencé les autres SC. En d'autres termes, on peut douter de l'existence d'une succession directe, lors du passage des termes entre SC-SPE et les autres SC. En effet, chaque SC semble plutôt montrer des dynamiques et des sélections qui leur sont propres, en fonction de choix liés sans doute aux genres, et aux besoins des groupes de locuteurs impliqués.

4.2.2. Indices distributionnels

L'analyse de l'environnement syntaxique de *boson de Higgs* fait apparaître des variations intéressantes. Nous avons examiné les structures *Adj boson de Higgs*, *boson de Higgs Adj*, *SN de boson de Higgs* dans chacun des 5 SC et étudié leur évolution comparée entre les deux périodes définies.

La Table 2 synthétise des résultats chiffrés pour le SC-SPE et pour le SC-PRESSE.

Table 2 - Synthèse de la distribution de *boson de Higgs* – SC-SPE et SC-PRESSE

	Uniquement dans SC-SPE avant 2011	Commun aux 2 périodes	Uniquement dans SC-SPE après 2011	Total
<i>N de boson de Higgs</i>	12	33	12	57
<i>Adj boson de Higgs</i>	0	0	2	2
<i>boson de Higgs Adj</i>	3	10	10	23
	15	43 (52,4 %)	24	82

	Uniquement dans SC-PRESSE avant 2011	Commun aux 2 périodes	Uniquement dans SC-PRESSE après 2011	Total
<i>N de boson de Higgs</i>	10	6	17	33
<i>Adj boson de Higgs</i>	9	3	6	18
<i>boson de Higgs</i>	0	1	5	6

<i>Adj</i>				
	19	10 (17,6 %)	28	57

Plusieurs phénomènes peuvent être notés :

- *boson de Higgs* s'intègre dans un environnement syntaxique plus varié dans SC-SPE que dans SC-PRESSE,
- *boson de Higgs* précédé d'un adjectif est très rare dans SC-SPE et cette structure n'apparaît que dans la deuxième période,
- le SC-SPE fait apparaître en plus grand nombre des éléments partagés par *boson de Higgs* dans les deux périodes par rapport à SC-PRESSE (52,4 % contre 17,6%).

Ces phénomènes peuvent s'expliquer relativement facilement. Tout d'abord, il n'est pas étonnant que le SC-SPE fasse apparaître un environnement syntaxique plus riche pour *boson de Higgs* que le SC-PRESSE. En effet, cette particule jouant un rôle majeur pour la physique des particules, il est normal qu'elle soit caractérisée par de nombreux modificateurs (*boson de Higgs* (*léger/pseudoscalaire/standard/supplémentaire/supersymétrique* ...)) ou qu'elle apparaisse en expansion dans beaucoup de syntagmes nominaux (*découverte/embranchement/désintégration/champ scalaire/masse...*) de (det) *boson de Higgs*). Dans certains cas, ces syntagmes nominaux sont d'ailleurs eux-mêmes des termes du domaine. On sait aussi que les nominalisations déverbiales (comme *découverte, désintégration*) sont plus fréquentes dans les corpus spécialisés que dans les corpus généraux (Condamines et Picton 2014), sans doute parce qu'elles viennent remplacer l'utilisation de verbes.

La présence d'adjectifs devant *boson de Higgs*, qui semble presque réservée au SC-PRESSE s'explique par le fait que presque tous les adjectifs dans cette position sont axiologiques et marquent une idée de sensationnalisme : *mystérieux, fameux, insaisissable, étrange, élusif*, etc. Ces appréciations contribuent

à attirer l'attention des lecteurs sur ce phénomène qu'est le boson de Higgs. Les deux seuls adjectifs présents devant *boson de Higgs* dans SC-SPE sont *seul* et *unique*.

Le fait que la presse fasse apparaître une instabilité distributionnelle plus importante entre les deux périodes que le SC-SPE peut s'expliquer par le fait que les journalistes adaptent leur discours, d'une part aux événements extérieurs (en l'occurrence la découverte du boson de Higgs) et, d'autre part, pour maintenir l'intérêt des lecteurs. Pour les scientifiques, la découverte du boson de Higgs est un événement certes important mais qui ne marque pas la fin de leur recherche. Un élément en faveur de cette différence de point de vue selon les groupes de locuteurs peut être confirmé par un autre phénomène. Ainsi, dans SC-PRESSE, on voit apparaître, avant 2011, beaucoup de phrases hypothétiques dans lesquelles *boson de Higgs* est utilisé :

- (1) *Le boson de Higgs, si le LHC démontre son existence, peut expliquer [...]*
- (2) *Le boson de Higgs pourrait être la clé de cette recherche*

Ces phrases hypothétiques disparaissent pratiquement dans la deuxième période.

En revanche, dans le SC-SPE, les phrases hypothétiques continuent à être utilisées, même après la découverte du boson de Higgs, parce que la « découverte » du boson de Higgs n'a pas épuisé les interrogations des scientifiques.

- (3) *Si la particule récemment découverte est bien le boson de Higgs*
- (4) *On suppose que l'unitarité de la théorie est restaurée par un boson de Higgs.*

Un autre aspect plaide en faveur d'une adaptation de la presse à des événements extérieurs. Si, dans la première période, la dimension sensationnelle est associée au boson de Higgs lui-même, dans la deuxième période, cette notion continue à apparaître mais associée, cette fois-ci, à la découverte elle-même.

- (5) Outre la découverte spectaculaire annoncée en juillet d'une particule aux caractéristiques compatibles avec le boson de Higgs
- (6) cette exposition immersive prend le parti de faire partager l'incroyable aventure humaine que représente la découverte du boson de Higgs.

Pour ce qui concerne les 3 autres SC, voici les chiffres concernant les contextes communs entre les deux périodes :

Table 3 - Synthèse de la distribution de boson de Higgs – SC-COM, SC-RAP et SC-VULG

	Uniquement dans SC-COM avant 2011	Commun aux 2 périodes	Uniquement dans SC-COM après 2011	Total
<i>Tous environnements</i>	5	6 16,2 %	26	37
	Uniquement dans SC-RAP avant 2011	Commun aux 2 périodes	Uniquement dans SC-RAP après 2011	Total
<i>Tous environnements</i>	12	9 20,9 %	22	43
	Uniquement dans SC-VULG avant 2011	Commun aux 2 périodes	Uniquement dans SC-VULG après 2011	Total
<i>Tous environnements</i>	19	9 20 %	17	45

Les chiffres se rapprochent de ceux du SC-PRESSE. Rappelons que les textes constituant ces corpus sont tous destinés à des non-experts. On peut donc faire l'hypothèse que l'évolution importante de la distribution de *boson de Higgs* (faible absence de contextes commun) est aussi liée à des événements extérieurs dans ces corpus.

5. Discussion et remarques conclusives

Les différentes observations menées dans notre étude permettent de dresser plusieurs constats quant au processus de déterminologisation, dans une perspective diachronique.

Du point de vue du fonctionnement des termes, les sous-corpus observés montrent des éléments communs importants, stables dans le temps, telles que des thématiques récurrentes, des changements de distribution semblables entre les périodes observées et des choix terminologiques communs. Ces « constantes » reflètent le processus de passage des termes et les liens évidents créés entre genres/degrés de spécialité, dans le temps. Néanmoins, nos résultats montrent également les spécificités, très nombreuses, des fonctionnements en jeu dans chacun des sous-corpus.

Les différents éléments observés mettent en effet tous en lumière une certaine indépendance des fonctionnements et des évolutions selon les sous-corpus. Ainsi, chacun opère une sélection propre des termes et de leur distribution, et révèle une évolution propre dans le temps. De ce point de vue, les termes susceptibles de se déterminologiser semblent également obéir à une sélection, possiblement en fonction de leur nature et des besoins de chacun des groupes de locuteurs.

Cet aspect questionne directement le rôle des genres impliqués dans le processus de déterminologisation. Plus précisément, une hypothèse souvent émise (plus ou moins explicitement) est celle d'un continuum entre textes spécialisés et presse. Nous avons développé ici cette réflexion en intégrant d'autres genres et degrés de spécialité impliqués dans ce processus (section 2.2). Or, nos observations en diachronie ne permettent pas toujours de mettre en évidence de réelles successions ou une « linéarité » nette entre les sous-corpus. Pour l'échantillon de termes observés ici en physique des particules, malgré certaines similarités (ou « constantes »), les sous-corpus observés semblent plutôt fonctionner de manière autonome. En poussant la réflexion un peu plus loin, le statut de la presse comme « genre d'arrivée » de ce continuum peut donc être discuté. Nos analyses tendent en effet plutôt à montrer que tous les genres et degrés de spécialité considérés ne sont pas des « étapes » du processus de déterminologisation, des textes spécialisés vers la presse, mais plutôt des « participants » de ce processus,

nécessairement liés, mais autonomes. Des questions similaires, sur la dynamique des discours et des genres sont posées par exemple par Moirand (2003) ou Beacco *et al.* (2002 : 278-279). Ces derniers, au sujet de la vulgarisation, expliquent :

Channels of knowledge transmission to non-specialists are probably less homogeneous than internal discursive forms, insofar as a scientific community has no control of knowledge dissemination.

Le processus de déterminologisation fait écho à cette complexité de dynamiques, et une réflexion sur la multimodalité des canaux par lesquels les termes passent et sur d'autres observables pertinents s'avère nécessaire.

Ces résultats peuvent aussi être rapprochés des travaux qui portent sur les communautés discursives et la diachronie. Dossena et Taavitsainen (2006 : 9) expliquent que les « [m]odes of presentation and styles of writing change in the course of time in response to changing needs of the discourse community ». Nos analyses illustrent par exemple la manière dont les « besoins » et intérêts des spécialistes évoluent différemment de ceux de la presse (et surtout de ses lecteurs). De même, Beacco (1995 : 38) signale :

[...] ainsi pourrait-on rendre compte, par exemple, de la variation diachronique de la production textuelle médiatique par rapport à la relative stabilité de discours normatifs comme ceux du droit ou de l'administration, fortement ancrés dans leur intertextualité antérieure.

Dans notre cas, les discours scientifiques peuvent sans doute se rapprocher de discours normatifs dans le sens où ils s'inscrivent dans une continuité (une intertextualité) forte, propre à l'évolution de la science et à la stabilité de ces communautés discursives.

L'intérêt de ces travaux réside également dans la place accordée à la diachronie dans l'analyse. Notre approche soutient cette idée, à travers l'exemple de la déterminologisation, et confirme l'importance de la diachronie pour dresser un portrait nuancé et complet des fonctionnements en jeu dans ce type de phénomène.

Du point de vue méthodologique, l'analyse diachronique n'est pas chose aisée⁵, mais reste possible (en particulier sur des diachronies courtes et contemporaines). Les possibilités offertes par la linguistique de corpus et les outils (dont la statistique) permettent d'explorer de nouvelles manières d'aborder les données diachroniques et de les analyser (Prévost 2015).

Dans notre étude, la mise en place d'un corpus « multi-comparable » a requis une attention particulière, qui permet d'alimenter cette réflexion, en particulier sur le découpage diachronique, encore peu questionné. Or, il nous semble clair que les dynamiques observées (et, partant, les conclusions qui sont dressées) peuvent diverger selon le découpage adopté. Une réflexion plus poussée sur les données diachroniques devrait donc être développée, pour les langues de spécialité.

Références

- ANTHONY Laurence (2018) : *AntConc (Version 3.5.7) [Computer Software]*. descriptif disponible sur <http://www.laurenceanthony.net/software>, Tokyo, Japan: Waseda University.
- BANKS David (2016) : « Diachronic Aspects of ESP », *ASp*, 69, p.97-111.
- BEACCO Jean-Claude (1995) : « À propos de la structuration des communautés discursives : beaux-arts et appréciatif », *Les Carnets du Cediscor*, 3, p. 136-153.
- BEACCO Jean-Claude, CLAUDEL Chantal, DOURY Marianne, PETIT Gérard et REBOUL-TOURÉ Sandrine (2002) : « Science in Media and Social Discourse: New Channels of Communication, New Linguistic Forms », *Discourse Studies*, 4(3), p. 277-300.
- CABRÉ M. Teresa (1998) : *La terminologie : Théorie, méthode et applications*, Ottawa, Les Presses de l'Université d'Ottawa et Armand Colin.

⁵ Ce que relève également la linguistique variationnelle. Völker (2002), par exemple, liste une série d'obstacles méthodologiques limitant les analyses diachroniques.

- CALBERG-CHALLOT Marie (2007) : « Quand un vocabulaire de spécialité emprunte au langage courant : Le nucléaire, étude de cas », dans John HUMBLEY (éd.), *Aspects de la recherche en langue de spécialité*, Université Paris VII, p. 71- 85.
- CONDAMINES Anne et PICTON Aurélie (2014) : « Étude du fonctionnement des nominalisations déverbales dans un contexte de déspecialisation », *Actes du Congrès Mondial de linguistique Française (CMLF)*, Berlin, 19-23 juillet 2014, p. 697-712.
- CONDAMINES Anne et PICTON Aurélie (à paraître) : « Textual Terminology: Origins, Principles and New Challenges ». In Marie-Claude L'HOMME et Pamela FABER : *Theoretical Approaches to Terminology*. Amsterdam/Philadelphia, John Benjamins.
- DELAVIGNE Valérie (2001) : *Les mots du nucléaire. Contribution socioterminologique à une analyse des discours de vulgarisation*. Thèse de Doctorat, Université de Rouen, France.
- DELAVIGNE Valérie (2003) : « Quand le terme entre en vulgarisation », *Actes de Terminologie et Intelligence artificielle, TIA*, p. 80- 91.
- DELAVIGNE Valérie (2020) : « De l'(in)constance du métalinguistique dans un corpus de vulgarisation médicale. », *Corela. Cognition, représentation, langage*, HS-31, <<http://journals.openedition.org/corela/11031>> (consulté le 15/10/2020).
- DOSSENA Marina et TAAVITSAINEN Irma (2006) : « From Synchrony to Diachrony » Dans Marina DOSSENA et Irma TAAVITSAINEN (eds), *Diachronic perspectives on Domain Specific English*. Bern, Peter Lang, p.7-17
- DROUIN Patrick (2003) : « Term Extraction Using Non-technical Corpora as a Point of Leverage », *Terminology*, 9(1), p. 99-117.
- DROUIN Patrick, MÉNARD Nathan et PAQUIN Annie (2006) : « Extraction semi-automatique des néologismes dans la terminologie du terrorisme », *Actes des 8^{èmes} Journées internationales d'Analyse statistique des Données Textuelles (JADT 2006)*, Besançon, France.
- DURY Pascaline (2008) : « The Rise of Carbon Neutral and Compensation Carbone: A Diachronic Investigation into the

- Migration of Vocabulary from the Language of Ecology to Newspaper Language and Vice Versa », *Terminology*, 14(2), p. 230–248.
- DURY Pascaline (2013) : « Que montre l'étude de la variation d'une terminologie dans le temps. Quelques pistes de réflexion appliquées au domaine médical », *Debate terminológico*, 9, p. 2-10.
- DURY Pascaline (2018) : *La dimension diachronique en anglais de spécialité : une approche terminologique*. Université Paris 7: note de synthèse pour l'Habilitation à Diriger des Recherches.
- DURY, Pascaline et LERVAD Susanne (2008) : « La variation synonymique dans la terminologie de l'énergie : approches synchronique et diachronique, deux études de cas. », *LSP and Professional Communication*, 8(2), p.66-79.
- DURY Pascaline et PICTON Aurélie (2008) : « Terminologie et diachronie : vers une réconciliation théorique et méthodologique ? », *Revue Française de Linguistique Appliquée*, 13(2), p. 31-41.
- FABER Pamela et LEÓN-ARAÚZ, Pilar (2016) : « Specialized Knowledge Representation and the Parameterization of Context », *Frontiers in Psychology*, 7(196), p. 1-20.
- FREIXA Judit (2006) : « Causes of Denominative Variation in Terminology - A Typology Proposal », *Terminology*, 12(1), p. 51-77.
- GAMBIER Yves (1991) : « Présupposés de la terminologie : Vers une remise en cause », *Cahiers de linguistique sociale*, 18, p. 31- 58.
- GAUDIN François (1993) : *Pour une socioterminologie : Des problèmes sémantiques aux pratiques institutionnelles*, Rouen, Publications de l'Université de Rouen.
- GAUDIN François (2003) : *Socioterminologie : Une approche sociolinguistique de la terminologie*, Paris, De Boeck-Duculot.
- GESMANN Markus et DE CASTILLO Diego (2011) : *Interface between R and the Google Visualisation API. GoogleVis package for R*.
- GRIES Stephan T., et HILPERT Martin (2008) : « The Identification of Stages in Diachronic Data: Variability-Based Neighbour Clustering. », *Corpora*, 3(1), p. 59-81.
- GRIES Stephan T. et HILPERT Martin (2012) : « Variability-based Neighbor Clustering: A Bottom-Up Approach to Periodization in Historical Linguistics. », Dans Terttu

- NEVALAINEN et Elizabeth TRAUGOTT (Eds.), *The Oxford handbook of the history of English* (pp. 134-144). Oxford: Oxford University Press.
- GUESPIN Louis (1991) : « La circulation terminologique et les rapports entre science, technique et production », *Cahiers de linguistique sociale*, 18, p. 59- 79.
- GUILBERT Louis (1965) : *La formation du vocabulaire de l'aviation*. Thèse de Doctorat, Université de Paris.
- GUILBERT Louis (1973) : « La spécificité du terme scientifique et technique », *Langue Française*, 17, p. 5- 17.
- GUILBERT Louis (1975) : *La créativité lexicale*, Paris, Larousse.
- HUMBERT-DROZ Julie (en cours) : « Circulation des termes entre langues de spécialité et langue générale : proposition d'un cadre théorique et méthodologique d'analyse du phénomène de la déterminologisation », Thèse de doctorat, Université de Genève/Université de Toulouse Jean-Jaurès.
- HUMBERT-DROZ Julie, PICTON Aurélie et CONDAMINES Anne (2019) : « How to Build a Corpus for a Tool-Based Approach to Determinologisation in the Field of Particle Physics », *Research in Corpus Linguistics*, (7), p. 1-17. doi:10.32714/ricl.07.01
- HUMBLEY John (2011) : « Vers une méthode de terminologie rétrospective. », *Langages*, 183(3), p. 51-62. doi:10.3917/lang.183.0051
- HUMBLEY John (2018) : *La néologie terminologique*, Limoges, éditions Lambert-Lucas, collection *La Lexicothèque*.
- JACQUES Marie-Paule (2005) : L'emploi de termes réduits comme révélateur de la centralité dans le domaine. *Actes des 7èmes journées scientifiques du Réseau "Lexicologie, Terminologie, Traduction"* (Agence Universitaire de la Francophonie) : « Mots, termes et contextes », Bruxelles, Belgique, p. 299-308.
- KILGARRIFF Adam, BAISA Vít, BUŠTA Jan, JAKUBÍČEK Miloš, KOVÁŘ Vojtěch, MICHELFEIT Jan, RYCHLÝ Pavel et SUCHOMEL Vít (2014) : « The Sketch Engine: ten years on », *Lexicography*, 1(1), p. 7-36.
- LEDOUBLE Hélène (2019) : « Vulgarisation scientifique et médiatisation de la science: Instabilité terminologique dans le domaine de la lutte biologique », *Terminology*, 25(1), p. 60-92. doi:https://doi.org/10.1075/term.00028.led
- LEDOUBLE, H. (2020) : « Term Circulation and Conceptual Instability in the Mediation of Science: Binary Framing of the Notions of

- Biological versus Chemical Pesticides », *Discourse & Communication*, 14(5), p.466–488.
- MEYER Ingrid (2000) : « Computer Words in Our Everyday Lives : How are they Interesting for Terminography and Lexicography? », *Actes de la conférence internationale EURALEX*, p. 39- 58.
- MEYER Ingrid et MACKINTOSH Kristen (2000) : « L' "étirement" du sens terminologique : Aperçu du phénomène de la déterminologisation », dans Henri BÉJOINT et Philippe THOIRON (Eds), *Le sens en terminologie*, Lyon, Les Presses Universitaires de Lyon, p. 198- 217.
- MEYER Ingrid, MACKINTOSH Kristen et VARANTOLA Krista (1998) : « From Virtual Sex to Virtual Dictionaries : On the Analysis and Description of a De-terminologized Word », *Actes de la conférence internationale EURALEX*, p. 645-654.
- MOIRAND Sophie (1999) : « Éléments de théorisation d'une linguistique du discours : L'exemple des discours sur les sciences dans les médias », *Modèles linguistiques*, 40, p. 5-21.
- MOIRAND Sophie (2003) : « Communicative and Cognitive Dimensions of Discourse on Science in the French Mass Media », *Discourse Studies*, 5(2), p. 175-206.
- MOIRAND Sophie (2007) : *Les discours de la presse quotidienne*, Paris, Presses Universitaires de France et Linguistique nouvelle.
- MØLLER Bernt (1998) : « À la recherche d'une terminochronie. » *Meta*, XLIII(3), p. 426-453.
- PICTON Aurélie (2014) : The Dynamics of Terminology in Short-Term Diachrony: A Proposal for a Corpus-based Methodology to observe Knowledge Evolution. Dans Rita TEMMERMAN et Marc VAN CAMPENHOUDT (Eds.), *The Dynamics of Culture-bound Terminology in Monolingual and Multilingual Communication, Terminology and Lexicography Research and Practice* (Vol. 16, pp. 159–182). Amsterdam/Philadelphie: John Benjamins.
- PICTON Aurélie (2018) : « Terminologie outillée et diachronie : éléments de réflexion autour d'une réconciliation », *ASp [En ligne]*, 74, p. 27-52.
- PICTON Aurélie, DROUIN Patrick et HUMBERT-DROZ Julie (à paraître) : Exploration et visualisation de la variation terminologique en corpus spécialisés complexes : réflexions et propositions méthodologiques. Dans Matthieu MANGEOT et Agnès TUTIN (Dir.), *Lexique(s) et genre(s) textuel(s)* :

- approches sur corpus*. Paris : Éditions des archives contemporaines, p. 99-116.
- PRÉVOST Sophie (2015) : « Diachronie du français et linguistique de corpus : une approche quantitative renouvelée. » *Langages*, Armand Colin (Larousse jusqu'en 2003), 197 (1), p. 23-45.
- R CORE TEAM (2018) : *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, <https://www-r-project.org> (consulté le 24/10/2018).
- RENOUF Antoinette (2017) : « Some Corpus-Based Observations on Determinologisation », *Neologica*, 11, p. 21-48.
- SIOUFFI Gilles, STEUCKARDT Agnès et WIONET Chantal (2012) : « Comment enquêter sur des diachronies courtes contemporaines ? », Actes du 3e Congrès Mondial de Linguistique Française, Lyon, France. pp.215-226.
- SWALES John (1990) : *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.
- UNGUREANU Ludmila (2006) : *L'interpénétration langue générale-langue spécialisée dans le discours d'internet*. Paris, Connaissances et Savoirs.
- URIELI Assaf (2013) : *Robust French syntax analysis: reconciling statistical methods and linguistic knowledge in the Talisman toolkit*. Thèse de doctorat, Université Toulouse Jean-Jaurès.
- VAN DER YEUGHT Michel (2009) : *Structurer l'anglais de la bourse : une approche des langues de spécialité par la spécialité*. Habilitation à diriger des recherches, Bordeaux : Université Victor-Segalen.
- VÖLKER, Harald (2002) : « Linguistique variationnelle historique et traitement de textes : présentation d'une méthode d'analyse », séminaire de l'ATILF, Nancy, [http://web.atilf.fr/IMG/pdf/seminaires/Seminaire_atilf_Volker_2002-11.pdf, dernier accès le 15.01.2021]

Aurélie PICTON
TIM-FTI, Université de Genève

Anne CONDAMINES

CLLE, CNRS et Université Toulouse Jean-Jaurès

Julie HUMBERT-DROZ
TIM-FTI, Université de Genève
CLLE, CNRS et Université Toulouse Jean-Jaurès