

CAHIERS DE L'ASSOCIATION INTERNATIONALE  
DES ÉTUDES FRANÇAISES

L'ÉCRITURE FRANCOPHONE  
EN BELGIQUE

SCARRON

LE PORTRAIT  
DANS LA CRITIQUE LITTÉRAIRE

Mai 2011

N° 63

Publié avec le concours du Centre National du Livre

Diffusion : LES BELLES LETTRES - 25, rue du Général Leclerc - 94270 LE KREMLIN-BICÊTRE



# LITTÉRATURE ET BASES DE DONNÉES : LE CAS DE LA BASE DE DONNÉES DU CIEL

*Communication de M. Björn-Olav DOZO*

(F.R.S. - FNRS - Université de Liège)

*au LXII<sup>e</sup> Congrès de l'Association, le 5 juillet 2010*

Lorsque l'on parle de bases de données en lettres, on pense souvent aux grands corpus littéraires numérisés, du type « Frantext » ou, pour la Belgique, « Beltext ». Ces bases de données sont surtout mobilisées par les linguistes. L'histoire de la littérature en fait peu d'usage (même si les choses évoluent, notamment avec les perspectives ouvertes par les travaux de Franco Moretti (1)). À côté de ces bases de données textuelles, l'histoire de la vie littéraire et la sociologie des écrivains a développé ses propres bases de données. C'est un exemple pionnier en la matière que je voudrais présenter aujourd'hui.

L'étude de la littérature belge a été menée au gré d'initiatives éparses jusqu'au début des années 2000, quand des projets de recherche à dimension collective se sont imposés. La base de données du CIEL est le fruit d'une initiative de recherche collective, collaboration entre l'Université Libre de Bruxelles et l'Université de Liège, dirigée par Paul Aron et Jean-Marie Klinkenberg, et coordonnée par Benoît Denis. Ce

(1) Moretti Franco, *Graphes, cartes, arbres. Modèles abstraits pour une autre histoire de la littérature*, Paris, Les Prairies ordinaires, 2008, « Penser/croiser ».

projet a été financé par la Communauté française de Belgique de 2002 à 2007 à Liège et à Bruxelles, puis de 2008 à 2013 à Bruxelles. Plusieurs thèses ont été réalisées en son sein (2).

La base a deux objectifs principaux : l'un patrimonial, l'autre plus « expérimental ». Le premier découle du constat que l'histoire de la littérature belge de langue française s'est écrite au gré d'initiatives éparses, provoquant de grandes disparités dans la documentation actuellement disponible, particulièrement en ce qui concerne les références bibliographiques. La constitution de la base de données vise d'abord à compiler les informations déjà existantes, à les corriger s'il y a lieu, puis à les mettre à jour. Le support informatique est le format idoine pour ce type de bibliographie en incessant devenir.

La base reprend, à l'heure actuelle, toutes les données prosopographiques disponibles dans les différents travaux existants, telles la *Biographie nationale* (3), la *Nouvelle Biographie nationale* (4), la *Bibliographie des écrivains francophones belges* (5) ou le *Dictionnaire des œuvres* (6). Ces données sont structurées et mises en relation de manière raisonnée. La base reprend des données factuelles, non analysées et le moins possible formatées. Il est en effet important, vu les différents objectifs

(2) Sur la genèse du projet CIEL, se reporter à l'article Paul Aron, Denis Benoît et Jean-Marie Klinkenberg, « Littérature belge et recherche collective », *Textyles*, n° 29, 2006, p. 90-97.

(3) H. Thiry, puis E. Bruylants et Ch. Christophe puis E. Bruylant (dir.), *Biographie nationale publiée par l'Académie Royale des Sciences, des Lettres et des Beaux-Arts de Belgique*, Bruxelles, H. Thiry – van Buggenhoudt, 1866-1986, 44 vol.

(4) *Nouvelle biographie nationale*, Bruxelles, Académie Royale des Sciences, des Lettres et des Beaux-Arts de Belgique, 1988, 6 vol.

(5) R. Brucher puis J. Detemmerman (dir.), *Bibliographie des Écrivains Français de Belgique*, Bruxelles, Palais des Académies, 1958-1988, 5 vol.

(6) V. Nachtergaele et R. Trousson (dir.), *Lettres françaises de Belgique. Dictionnaire des œuvres. I. Le roman*, Louvain-la-Neuve, Duculot, 1988. Ch. Berg et R. Frickx (dir.), *Lettres françaises de Belgique. Dictionnaire des œuvres. II. La poésie*, Louvain-la-Neuve, Duculot, 1988. M. De Grève, J. M. d'Heur et R. Pouilliart (dir.), *Lettres françaises de Belgique. Dictionnaire des œuvres. III. Le Théâtre et l'Essai*, Louvain-la-Neuve, Duculot, 1989. R. Frickx et alii, *Lettres françaises de Belgique. Dictionnaire des œuvres. IV. 1981-1990*, Louvain-la-Neuve, Duculot, 1994.

de la base, de ne pas uniformiser l'information récoltée dans les différentes sources sous des catégories qui ne seraient adaptées qu'à un seul type d'exploitation. Par exemple, pour les professions, si l'on adopte une nomenclature socio-professionnelle dans la base de données, on perd les spécificités onomastiques mises en avant dans les sources pour les désigner, qui ont pourtant un intérêt patrimonial, mais aussi scientifique si l'on travaille sur ces noms de professions. La sélection des informations dans les sources se concentre uniquement sur les éléments factuels, mais ne normalise pas les étiquettes lors de l'encodage. Nous avons seulement mis au point des catégories très vastes qui offrent une première structure thématique (les professions, les activités littéraires et culturelles) et descriptive (nom de l'activité, lieu, qualité de la personne qui exerce cette activité) des données, afin d'éviter le piège d'une base de données en texte plein qu'il est impossible d'interroger selon des critères définis.

Ce départ entre la structuration des données et le codage est un point important, qui devrait permettre de ne pas cloisonner la base du CIEL dans un type d'usage unique. En ce qui concerne notre exploitation particulière, nous présenterons quelques exemples de résultats fondés sur un traitement statistique des données. Mais avant cela, il nous semble intéressant de présenter les grandes lignes de la base de données du CIEL.

Nous nous contenterons de mettre en évidence la philosophie qui a présidé à l'établissement des structures de cette base. Pour le détail de son fonctionnement, la liste des conventions et les choix opérés dans le classement d'informations ponctuelles, on se référera utilement au mode d'emploi de la base (7) et aux articles de présentation et de problématisation du projet (8).

(7) Voir le site du projet, <http://www.ciel-litterature.be>.

(8) Dozo Björn-Olav et Fréché Bibiane, « Réseaux et bases de données », [in] Daphné de Marneffe et Denis Benoît (dir.), *Les Réseaux littéraires*, Bruxelles, Le Cri - ULB - ULg, 2006, p. 86-108 ; Paul Aron, Benoît Denis et Jean-Marie Klinkenberg, *op. cit.*

## LA SECTION « AUTEURS »

La section « auteurs » reprend des informations factuelles socio-historiques sur les écrivains.

L'unité de cette partie de la base CIEL est l'écrivain belge de langue française, c'est-à-dire la personne physique qui a eu la nationalité belge à un moment de sa vie et une activité littéraire (9) en langue française entre 1920 et 1960 (10). Chaque personne est encodée sous son nom d'écrivain, qui peut être différent de son nom d'état civil et de pseudonymes qu'il utilise occasionnellement : il s'agit du nom par lequel l'écrivain signe habituellement ses œuvres. Sa fiche, à laquelle correspond un numéro unique, reprend aussi son nom d'état civil et ses pseudonymes : il n'y a ainsi qu'une fiche par personne physique. On désignera dorénavant cette personne physique par le syntagme « agent », terme sociologique qui vise à neutraliser les effets impliqués par des mots comme « auteur », « écrivain », etc.

La base de données du CIEL est une base de données relationnelle, c'est-à-dire qu'il s'agit d'une base de données structurée suivant les principes de l'algèbre relationnel. Celle-ci est un ensemble d'opérations formelles sur les relations existant entre différents éléments, qui sont répartis en plusieurs tables, selon leur nature. Il existe ainsi des relations entre un élément d'une table et plusieurs éléments d'une autre table – ces relations sont désignées comme des

(9) L'activité littéraire est définie en creux, dans un but pratique de constitution de la base de données : par activité littéraire, on entend, en première approximation, la publication en volume d'au moins une œuvre littéraire, c'est-à-dire qui n'est ni scientifique, ni technique, ni juridique, ou la publication d'une contribution dans une revue à dominance littéraire (c'est-à-dire pas purement informative, ni spécialisée dans d'autres domaines que la littérature), voire seulement la participation à la vie littéraire de l'époque, que cela soit par l'animation ou la direction d'une revue littéraire, ou la direction d'une maison d'édition belge, etc.

(10) Le corpus actuel est un corpus de départ. La base sera ouverte à tous les auteurs belges par la suite.

relations de un à plusieurs – ou entre plusieurs éléments d’une table et plusieurs éléments d’une autre table – relations dites de plusieurs à plusieurs. L’algèbre relationnelle repose sur la combinaison de ces types de relation. Dans le cas de la base « auteurs », l’élément principal est l’écrivain, auquel sont reliés différents éléments secondaires : ses résidences, sa famille, ses professions, ses formations scolaires, son cursus social, sa sociabilité littéraire, ses activités littéraires et culturelles, ses marques de reconnaissance dans le monde littéraire, ses conférences, etc.

Comme le montre le diagramme suivant, l’élément principal est occupé par les renseignements d’état civil qui permettent d’identifier l’écrivain de manière univoque (nom, prénom, pseudonyme(s), date de naissance, date de mort, nationalité, etc.). Sont liés à ce centre les renseignements annexes, portant sur la famille, les résidences, les professions, les activités littéraires et culturelles, etc.

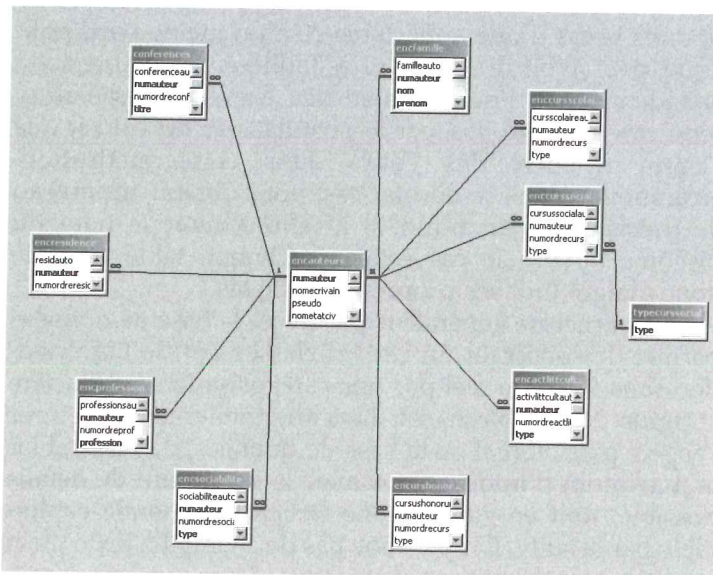


Figure 1 – Structure de la section « auteurs »

Les tables qui permettent de structurer les renseignements annexes sont conçues sur un même modèle : chaque table (résidences, professions, formations scolaires, cursus social, etc.) comprend trois grands éléments : le type, la dénomination de ce type et la fonction exercée par l'agent. Chaque liste de types varie en fonction de la table. Par exemple, pour la table « formations scolaires », les différents types sont : primaires, secondaires, secondaires inférieures, secondaires supérieures, école normale, supérieures non universitaires, universitaires, artistiques, autres. Vient ensuite la dénomination du lieu ou de l'institution relatif au type sélectionné (par exemple le nom de l'école dans la table des formations scolaires, ou encore le nom du parti politique auquel adhère l'agent dans la table du cursus social). Enfin vient la fonction occupée par l'agent dans cette institution ou ce lieu : pour prendre un autre exemple, dans la table du cursus social, la fonction de l'agent est présente, car celui-ci, dans le cas d'une appartenance à un même parti politique, peut avoir différents rôles (adhérent, membre actif, président, etc.). Pour la table des formations scolaires, cette case n'existe pas, car le postulat par défaut est que l'agent a suivi des cours dans cette institution. Néanmoins, dans ce dernier cas, pour obtenir un niveau de précision assez grand, il a fallu ajouter le nom du diplôme obtenu le cas échéant. Chaque table possède donc malgré tout ses propres spécificités.

Cette structure générale tripartite de la base de données permet de conserver un certain classement de l'information sans forcer le réel par une catégorisation outrancière et rigide. Notre objectif est, dans un premier temps et pour l'aspect patrimonial de la base de données, de rassembler le maximum d'informations avec le maximum de détails possible, tout en gardant une structure générale exploitable par la suite. Il n'y a donc pas de champ en texte plein et continu, sans balisage ni structure. Dans le même esprit, et toujours par choix raisonné, nous n'avons pas opté pour



des listes fermées de choix pour chaque champ : le dépouillement des sources s'apparente à une démarche exploratoire avant tout, où l'on découvre pour quasiment chaque auteur des réalités auxquelles nous n'avions pas songé et qui, si nous avons choisi des listes fermées, ne seraient pas entrées dans nos catégories... Seuls les types de chaque table sont des listes fermées, qui permettent d'opérer un premier tri non contraignant dans les données récoltées.

Ces trois éléments récurrents (type, dénomination et fonction) sont complétés par des renseignements chronologiques (date de début et date de fin), et par des données nécessaires en propre à la table (par exemple – nous l'avons déjà cité – le diplôme obtenu pour la table des formations scolaires).

Reste que ce qui différencie un écrivain d'un non-écrivain, ce sont ses œuvres. L'un des critères pour entrer dans la base du CIEL, comme on l'a vu, est d'avoir publié au moins un texte dans une revue littéraire ou une œuvre en volume. Commençons par ce dernier cas et voyons quelles informations sont collectées dans la base « œuvres ».

## LA SECTION « ŒUVRES »

Techniquement, la section « œuvres » n'est qu'une table supplémentaire de la base « auteurs », au même titre que la table des professions par exemple. Mais les œuvres sont plus qu'un type de renseignements parmi d'autres : c'est par elles que l'écrivain est écrivain. On a donc pris l'habitude de considérer la section « œuvres » comme une entité à part, tant dans son mode d'emploi que dans l'interface de la base de données du CIEL : à une interface qui lui est propre correspond un chapitre spécifique expliquant son fonctionnement.

Les « œuvres » sont une réalité littéraire que chacun manie quotidiennement. Pourtant, leur définition pose de nombreux problèmes, ne serait-ce que par l'abondan-

ce de celles-ci. Il s'agit là d'une question très générale, qui dépasse la confrontation entre l'informatique et le littéraire. Le concept d'« œuvre » est utilisé de manière fourre-tout, pour parler d'un texte, de ses éditions, des livres qui le contiennent, de ses adaptations, de ses traductions, etc. Un flottement terminologique et essentiel règne autour de ce terme. Cette question pourrait faire l'objet de nombreux débats mais nous nous contenterons de l'aborder sous l'angle pragmatique de la base de données, en délimitant précisément l'objet qui est encodé dans la base.

Pour ce faire, nous devons élargir notre problématique et analyser les répertoires de productions culturelles. Ces répertoires et catalogues ont été abondamment étudiés par l'*International Federation of Libraries Association* (IFLA), organisation indépendante, internationale, non lucrative et non gouvernementale, qui se donne pour but d'encourager la standardisation de services bibliothécaires de haute qualité et qui regroupe les plus grandes bibliothèques du monde.

L'IFLA a créé un modèle de catalogage de ces productions culturelles, modèle utilisé dans les répertoires des lieux de conservation (bibliothèques, médiathèques, etc.). Il présente les réalités culturelles au travers de leurs différentes incarnations et met en évidence une hiérarchisation de ces incarnations, chacune d'elles se situant à un degré d'abstraction différent.

Ce modèle a été adapté spécifiquement pour nos besoins. Le schéma suivant explique les différents niveaux d'incarnation de l'objet littéraire, avec pour chaque niveau les différents intervenants sur le texte.

	<b>Incarnation de l'objet littéraire</b>	<b>Intervenants sur le texte</b>
<b>Niveau 1</b>	travail ou œuvre	auteur
<b>Niveau 2</b>	expression de l'œuvre	niveau 1 + traducteur, adaptateur
<b>Niveau 3</b>	manifestation de l'œuvre : édition	niveau 2 + préfacier, postfacier, commentateur, éditeur scientifique, etc.
<b>Niveau 4</b>	item ; incarnation de l'œuvre, objet papier	niveau 3 + propriétaire

Cette structure pyramidale permet de différencier plusieurs réalités : le « travail » peut s'assimiler à l'« œuvre » (niveau 1), définie par un titre général et un nom d'auteur, « premier » à avoir travaillé et créé cette œuvre. Cette œuvre connaît alors plusieurs « expressions » (niveau 2) : on retrouve pour ce deuxième niveau d'abstraction les mêmes informations (à savoir l'auteur et le titre), plus la langue de l'expression et sa forme générique (adaptation au théâtre par exemple). Ces « expressions » de l'œuvre prennent forme dans différentes « éditions » (niveau 3). Ce troisième niveau d'abstraction comprend les informations indispensables à la description complète de l'édition d'une œuvre. Ces éditions elles-mêmes s'incarnent dans de multiples « objets-livres » (niveau 4) (les volumes que l'on peut tenir en main), qui peuvent avoir leurs particularités propres (dédicace personnalisée par exemple).

Un catalogue de bibliothèque utilise le niveau 4 pour référencer les livres. Notre but n'est pas de référencer des livres mais de produire des statistiques sur des auteurs, des éditeurs, etc., et d'établir des bibliographies par auteur. Nous avons choisi d'utiliser, dans la table principale, le troisième niveau, celui de l'édition. Ainsi, la table « œuvres » recense l'ensemble des ouvrages publiés par un auteur donné – en ce compris les traductions qu'il a

faites lui-même de son œuvre ou de celle d'autres écrivains. Pour chaque édition d'ouvrage, la base de données indique les éventuels auteurs, traducteurs, adaptateurs, collaborateurs, préfaciers, *postfaciers*, commentateurs, illustrateurs, etc. Elle spécifie s'ils sont belges afin qu'un lien puisse automatiquement être fait avec une fiche dans la base « auteurs » si c'est le cas. Elle donne aussi diverses informations sur le livre pris dans la matérialité de son édition : format, pagination, etc.

La table des traductions et des adaptations (au cinéma, à la télévision, à l'opéra, en bande dessinée, sous forme de chanson ou peinture, etc.) d'œuvres belges se situe quant à elle au deuxième niveau, c'est-à-dire au niveau de l'expression de l'œuvre. En effet, à ce stade, nous désirons uniquement dénombrer les adaptations et traductions d'une œuvre belge donnée, voire connaître la langue de sa traduction ou le genre de son adaptation, mais non trouver le nombre exact des éditions traduites ou adaptées.

Une fois explicitée l'unité retenue pour notre base de données (l'édition), il nous reste à énoncer les critères qui nous permettent d'y intégrer une édition d'œuvre. Ces critères sont intimement liés à la nationalité des intervenants sur cette œuvre. Par convention, si l'un de ces intervenants – qu'il soit auteur, traducteur, adaptateur, préfacier, commentateur, éditeur scientifique, etc. – est de nationalité belge et se trouve donc dans la base « auteurs », l'ouvrage auquel il prend part est intégré dans la table « œuvres ». Cette convention se justifie très pragmatiquement : il est dans nos objectifs de tendre vers la complétude des références relatives aux auteurs belges, et ce moyen a semblé le plus économique pour y parvenir sans que la gestion de la base ne devienne impossible (11).

(11) Si un auteur belge est traduit par un auteur belge, la fiche doit apparaître dans les deux tables, afin que la démarche établie reste cohérente. La première table doit en effet contenir toutes les éditions des œuvres de l'auteur belge, mais aussi toutes les éditions des œuvres du traducteur belge (dont chaque traduction est une œuvre). La deuxième, quant à elle, doit contenir toutes les adaptations et traductions des œuvres de l'auteur belge. Néanmoins, suivant

La base « œuvres » fournit également un résumé de chaque œuvre, à l'instar du *Dictionnaire des œuvres* cité plus haut. Elle se propose donc de devenir un *Dictionnaire électronique des œuvres littéraires belges de langue française*.

## LA SECTION « REVUES »

Techniquement et conceptuellement, la base « revues » est différente de la base « auteurs/œuvres ». Elle présente une structure enchâssée à plusieurs étages. Avant d'examiner en détail sa composition, il faut se pencher sur sa structure générale et sur la terminologie employée. L'objet qui nous a servi d'unité est l'article de revue, conventionnellement appelé « contribution ». On situe les contributions par rapport à la « livraison » (en langage courant, le « numéro ») de revue dans laquelle elles paraissent. L'ensemble des livraisons constitue la « revue », qui est elle-même caractérisée entre autres par un « état » – nous allons revenir sur cette étiquette et ce qu'elle désigne. Dans une fiche particulière, on caractérise par ailleurs le contenu de la « contribution », notamment en examinant les « sujets traités dans la contribution ».

Il faut donc définir ce que l'on entend par « état », « revue », « livraison », « contribution » et « sujets traités dans la contribution ». Ces termes renvoient à des réalités hiérarchisées entre elles : « revue » > « états » > « livraisons » > « contributions » > « sujets traités dans la contribution ». En effet, une « revue » peut se présenter sous différents « états », chacun de ceux-ci étant valable pour une ou plusieurs « livraisons » ; au sein de chaque livraison, on

la différence de niveau entre les deux tables, les informations encodées pour la fiche de la traduction de la première table et celle de la deuxième table ne seront pas identiques. La première sera plus complète que la deuxième, vu qu'elle renseigne une édition, tandis que l'autre ne signalera que l'existence d'une traduction, sans plus de précisions. Il est même possible que la première table contienne plusieurs fiches (une par édition de la traduction) tandis que la deuxième ne contiendra qu'une fiche (celle signalant la traduction).

identifie les « contributions » et, en dernière instance, les « sujets traités dans la contribution ».

Une revue littéraire est un périodique : elle est publiée sous un titre généralement récurrent et a au moins un rapport avec la littérature (soit qu'elle en publie, soit qu'elle abrite une rubrique critique qui lui est consacrée). Elle est identifiée dans la base uniquement par un numéro, qui la situe dans une liste constituée à partir du répertoire de Paul Aron et Pierre-Yves Soucy (12).

L'état est un ensemble de données récurrentes de livraison en livraison (titre, sous-titre, adresse de rédaction, format, tirage, conditions d'abonnement, etc.). Tant que ces données restent identiques, les livraisons renvoient au même état. Dès qu'une caractéristique subit une modification (changement d'adresse, de sous-titre, de format, etc.), on crée un nouvel état. Au cours de sa publication, chaque revue est susceptible de se manifester sous plusieurs états.

La livraison est constituée des feuilles de papier que l'on peut avoir en main. Il s'agit donc d'un objet matériel, un « numéro » de la revue dans le langage courant.

La contribution est un terme générique qui permet de rassembler sous le même vocable tous les textes publiés dans une même livraison de la revue, qu'il s'agisse d'un poème, d'une nouvelle, d'un extrait de roman, d'une critique ou encore d'une chronique (à l'exception des annonces publicitaires et des offres). Une contribution est une portion homogène du continuum textuel de la revue, distinguée des autres par une signature unique. Elle signifie la présence textuelle continue d'un auteur.

On obtient, à partir de ces définitions, un schéma d'ensembles contenus les uns dans les autres :

(12) Paul Aron et Pierre-Yves Soucy, *Les Revues littéraires belges de langue française, de 1830 à nos jours*, Bruxelles, Labor, 1998.

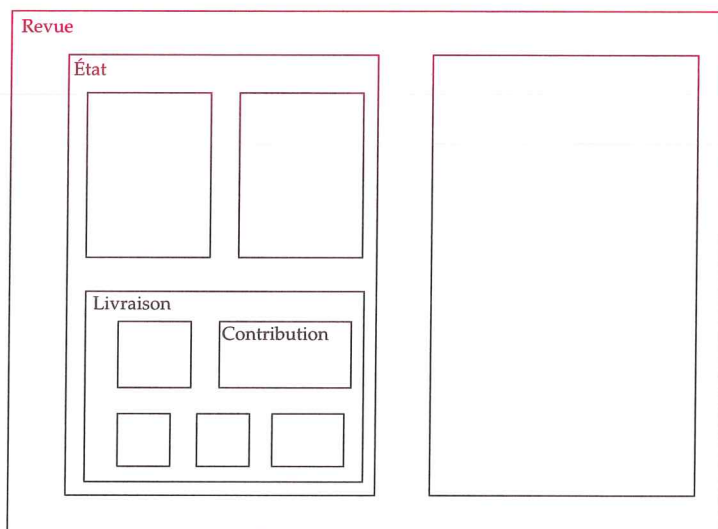


Figure 2 – Structure conceptuelle de la section « revues »

Pour une raison pratique (liée aux impératifs matériels d'encodage, qui se fait à partir des « livraisons »), la structure de la base ne reflète qu'en partie cette hiérarchisation conceptuelle. En effet, il n'est bien souvent pas possible de disposer de la revue dans son intégralité, et ainsi de fournir une description exhaustive des états lors de la première phase de saisie des données, à partir des quelques livraisons dont on dispose au départ. La conséquence principale est l'absence de liaison entre les tables « revues » et « états », et l'existence de relations non caractérisées entre les tables « états » et « livraisons ». Cela induit un contrôle moins automatisé du contenu de la base, particulièrement dans la table « états ». Mais cet état transitoire de la base de données pourra être corrigé lors d'un stade ultérieur de développement, si la masse d'informations concernant les états des revues permet une description exhaustive.

Il existera plusieurs relations entre les tables de la section « revues » et les autres tables (« auteurs » et « œuvres »). Ainsi, pour la table des « sujets traités dans la contribution »,

un lien est prévu vers la fiche de l'auteur, de l'œuvre ou de la revue si le sujet traité renvoie à un objet référencé dans la base. De même, chaque auteur d'une contribution sera relié à sa fiche dans la base « auteurs », s'il répond aux conditions pour y entrer. Il en va de même pour les illustrateurs si ceux-ci sont aussi écrivains, pour les conférenciers repris dans la table « annonces », ou encore pour les responsables de dossiers dans certaines livraisons. Enfin, l'organigramme du comité de rédaction d'une revue renvoie pour chaque membre à sa fiche dans la base « auteurs ». On le voit, les interconnexions entre les différentes tables sont au cœur même de l'organisation de la base de données du CIEL. Ces interconnexions sont la réelle force de notre base, tant pour l'exploitation statistique des données que pour la consultation de ces données via un site web destiné au grand public.

## USAGES SCIENTIFIQUES ET GRAND PUBLIC DE LA BASE CIEL

La structure de la base de données du CIEL met l'accent, comme on a pu s'en rendre compte, sur les liens qu'entretiennent différents objets, l'un apportant de l'information concernant l'autre. Cette non-hiérarchisation apparente offre une grande souplesse dans l'organisation de l'information, et ne privilégie pas l'étude future d'un objet par rapport à un autre. Les interrogations possibles sont quasiment infinies : elles dépendent du chercheur, de ses intérêts et des contraintes structurelles de la base que l'on vient d'exposer.

Si la base peut servir de vaste réservoir à informations ponctuelles, sa spécificité est de permettre le traitement sériel des données. Nous avons ainsi pu réaliser différentes études concernant les professions des écrivains (13),

(13) Dozo Björn-Olav, « La présence des juristes dans l'institution littéraire belge. Examen de l'évolution des trajectoires scolaires et professionnelles des écrivains dans l'entre-deux-guerres », *Textyles. Droit et littérature*, n° 31, 2007, p. 28-46, et Dozo Björn-Olav, « Portrait statistique de l'écrivain journaliste en Belgique francophone entre 1920 et 1960 », *Textyles*, n° 39, 2010, à paraître.



les genres qu'ils pratiquaient et les éditeurs qui les publiaient (14), les lieux de sociabilité qu'ils fréquentaient (15) ou encore le réseau formé par les préfaces dans l'entre-deux-guerres (16).

Ces différentes études se trouvaient facilitées par la structure même de la base de données, qui comprend, comme on l'a vu, de nombreuses catégories. Néanmoins, certaines interrogations ne trouvent pas une formalisation directe au sein de la base : certaines catégories n'existent pas en tant que telles. Cela ne signifie pourtant pas que la base de données est inutile dans ce cas. Prenons un exemple concret : dans la base « revues », il n'a pas été prévu de champ pour reprendre le nom des traducteurs de contribution d'une langue étrangère en français. Pourtant, il reste possible d'établir une liste des textes traduits en faisant une recherche sur l'expression « \*trad\* » dans le champ « commentaire » de chaque contribution. L'idée était de perdre le moins d'informations possible : ce qui ne pouvait être formalisé au sein de champs idoines pouvait trouver une place dans les champs de commentaires. Tout le défi était donc de trouver un équilibre entre la création de nouvelles catégories au fur et à mesure des dépouillements (afin que tout ne prenne pas place en commentaires) et la surabondance de catégories trop spécifiques et donc, au final, généralement vides. Cet équilibre se retrouve dans les interrogations possibles : la structure de la base de données conditionne les requêtes, mais elle n'en établit pas une liste fermée dont on ne peut sortir. Il reste au chercheur un vaste « espace des interrogations possibles ».

(14) Dozo Björn-Olav, « La production littéraire des Belges francophones durant l'entre-deux-guerres. Examen d'un premier inventaire », *Mémoires du livre / Studies in Book Culture*, Vol. 1, n° 2, 2010, à paraître.

(15) Dozo Björn-Olav, « Sociabilités et réseaux littéraires au sein du sous-champ belge francophone de l'entre-deux-guerres », *Histoire et Mesure*, n° XXIV-1, 2009, p. 43-72.

(16) Dozo Björn-Olav, « Structure de l'espace relationnel des auteurs francophones belges de l'entre-deux-guerres », dans Marie-Pier Luneau et Josée Vincent, *La Fabrication de l'auteur*, Québec, Nota Bene, 2010, p. 183-203, à paraître.

Le principe de la base est ainsi que le chercheur dispose d'une source d'informations structurées, vérifiées (17) et centralisées, qu'il peut enrichir au fil de ses recherches.

La base du CIEL est également un outil pour le grand public ou les enseignants de littérature : grâce au site du groupe de recherche (18), une version publique de la base de données est disponible pour consultation gratuite via internet. Le lecteur intéressé y trouvera de courtes notices biographiques, des bibliographies, des résumés d'œuvres et des sommaires de revues. Comme toutes les tables sont interconnectées, il est possible d'avoir pour un écrivain la liste de ses articles en revues (19), les lieux d'édition où il a publié ou encore les auteurs qui l'ont préfacé. Tout cela s'obtient au moyen d'un moteur de recherches qui autorise de multiples croisements.

Le site public est aussi un moyen de fédérer une communauté autour de cet objet : il est en effet possible de suggérer des ajouts ou des modifications à la base. Tout apport d'informations originales est une contribution précieuse à ce vaste « travail en cours » collectif.

Björn-Olav DOZO

(17) Les sources utilisées pour le moment sont celles considérées comme les plus sûres ; elles ont néanmoins fait l'objet de recoupements entre elles et, quand les données ne concordaient pas, d'un retour à la source originale. Le dépouillement continue à se faire par sources (et non par thème par exemple), et chaque information doit comporter son origine.

(18) <http://www.ciel-litterature.be>, site sur lequel la version publique de la base de données est en libre accès, moyennant une inscription gratuite préalable.

(19) Le dépouillement est encore en cours et le nombre de revues traitées est encore modeste, mais il progresse régulièrement.