



CRÉATION ET UTILISATION DE BASES DE DONNÉES POUR L'ENSEIGNEMENT DE LA THÉORIE DU RÉCIT ET DE LA SÉMIOTIQUE

Romain Gaudreault, chercheur indépendant
romain.gaudreault@videotron.ca

Résumé

Dans les ateliers de composition littéraire à l'école, les bases de données vraies ou fictives, sur support papier ou électronique, servent de mémoires auxiliaires externes, allègent la recherche de matériaux pour l'invention d'histoires et permettent de renouveler les méthodes d'enseignement du récit et de la sémiotique, entendue comme la science des signes et de leur utilisation dans la production et la communication de messages, y compris des récits. En amont, une telle approche renouvelle automatiquement les méthodes de recherche des spécialistes en études littéraires qui sont aussi des sémioticiens.

Introduction

En 1967, en plein cœur du mouvement de renouveau des études littéraires lancé en France par les Greimas, Genette et compagnie, l'écrivain italien Italo Calvino prononce une conférence sur un thème tout à fait curieux et, en apparence, éloigné des préoccupations des théoriciens français du récit : les machines littéraires. Pour cet auteur, « l'acte d'écrire n'est qu'un processus combinatoire entre des éléments donnés » (Calvino, 1984 [1967] : 21). Il n'est donc pas surprenant qu'il envisage la création d'une machine ou d'un automate littéraire qui explorerait toutes les combinaisons possibles de matériaux déterminés et proposerait une série d'histoires entre lesquelles le lecteur n'aurait plus qu'à choisir. Certains lecteurs pourraient en effet trouver amusant et pratique qu'une machine réalise toutes les étapes de la composition littéraire, jusqu'à et y compris la production d'un texte dans la langue de l'utilisateur. Cependant, envisageons la question du point de vue des rédacteurs, ou des scripteurs, comme on appelle les jeunes des écoles qui participent à des ateliers de composition littéraire. Ces scripteurs pourraient-ils compter au moins sur un outil qui faciliterait le processus combinatoire ? L'informatique offre un tel outil.

Dans le présent travail, nous allons décrire une approche de l'analyse et de la production de récits de fiction qui renouvelle à la fois les méthodes de recherche en études

littéraires et les méthodes d'enseignement de la théorie du récit : les bases de données. Dans une première partie, nous exposerons les principaux facteurs à prendre en compte dans la création d'une base de données pour l'invention d'histoires et, dans la seconde, nous montrerons comment utiliser ces bases de données. Soulignons que la présente étude relève de la sémiotique, entendue comme la science des signes et de leur utilisation dans la production et la communication de messages, y compris des récits.

1. Création d'une base de données

Dans cette première partie, après avoir dit ce qu'il faut entendre d'une façon générale par « base de données », nous examinerons successivement les divers degrés de validité des données, les domaines d'application des bases de données, les catégories qui servent à regrouper les données et les étapes de l'élaboration d'une base de données pour l'invention d'histoires.

1.1 Qu'est-ce qu'une base de données ?

Selon un dictionnaire d'informatique courant, une base de données est un « regroupement, dans un ensemble structuré, des données d'un ou plusieurs domaines d'application, afin de rendre ces données accessibles en permanence à l'ensemble des programmes utilisateurs » (Morvan, 1996 : 29). De cette définition, nous retiendrons d'abord qu'une base de données a un domaine d'application et que l'ensemble des données est « structuré ». Cela apparente la base de données à ce que Moles appelle un « répertoire », comme un répertoire d'œuvres symphoniques classiques, de signes typographiques ou de mots (Moles, 1972 [1958] : 50, 190). Ensuite, nous notons que la base de données est accessible par l'intermédiaire d'un programme d'ordinateur. Dans le contexte du présent travail, nous dirons qu'une base de données a un domaine d'application, qu'elle contient des données regroupées en sous-ensembles que nous appellerons « catégories » et qu'elle peut aussi être utilisée manuellement, sans ordinateur.

1.2 Validité des données

Il y a deux types de jugements de validité ou de vérité des données : en premier lieu, le jugement de réalité, qui porte sur l'existence ou la non-existence de la réalité désignée par les données ; en second lieu, le jugement de valeur, qui porte sur l'importance ou la non-importance de la réalité désignée par les données pour celui qui les recueille.

Pour les besoins du fonctionnement d'une société commerciale, on ne compte que deux degrés de validité : une donnée est normalement vraie, mais il arrive qu'elle soit fausse, c'est-à-dire qu'elle ne corresponde pas ou ne corresponde plus à la réalité. La vérité d'une

donnée n'est jamais affirmée comme telle. Elle fait partie des attentes du destinataire de la donnée en vertu d'un « principe de coopération » élémentaire qui, selon Grice (1989 : 26-28), fonde les échanges de messages dans toute société. Quant à la fausseté d'une donnée, elle peut être affirmée après coup par l'auteur de la donnée lui-même ou par une autre personne. La fausseté est due soit à une erreur de l'auteur de la donnée, soit à une manœuvre délibérée de sa part.

Outre les données vraies et les données fausses, on peut concevoir un troisième degré de validité, les données fictives. Par « donnée fictive », il faut entendre une donnée présentée dès le départ comme fausse. Ce jugement initial explicite annule le jugement de vérité normalement sous-entendu dans tout échange de messages. L'affirmation qu'un message est une fiction est fréquente chez les auteurs d'œuvres littéraires : dès que le terme « roman » figure sur la page couverture d'un livre, le lecteur est prévenu : toute l'histoire est fondée sur des données fictives, c'est-à-dire imaginées, inventées par l'auteur. Autrement dit, le terme « roman » inscrit sur la couverture est un jugement de réalité porté par l'auteur sur le contenu du texte qui va suivre. Lejeune (1975 : 27) appelle « pacte romanesque » cette déclaration initiale de l'auteur d'un roman. Ce pacte est parfois rompu. Lejeune (1975 : 31) donne l'exemple du récit intitulé *Le Sabbat*, de Maurice Sachs, et accompagné du sous-titre « roman » lors de sa parution en 1960. Or ce récit avait déjà été publié chez un autre éditeur en 1946 avec le sous-titre *Souvenirs d'une jeunesse orange* ! Voilà un cas de fausse fiction qui s'ajoute aux trois autres degrés de validité mentionnés plus haut : la vérité, la fausseté et la fiction.

Tout message, qu'il soit long, de type « article de journal » ou « roman », ou bref, comme une donnée, est accompagné dès le départ, avons-nous dit, d'un jugement quant à sa vérité (implicitement, du fait du principe de coopération de Grice) ou quant à sa fausseté (explicitement, toujours du fait du principe de coopération). Ce jugement de l'émetteur est catégorique : selon l'émetteur, le destinataire peut être certain que ce message est vrai (cas ordinaire), ou qu'il est faux (cas de la fiction). Cependant, le jugement de l'émetteur peut être hypothétique, c'est-à-dire que l'émetteur peut ne pas être certain que le message est vrai ou qu'il est faux. Cette incertitude complexifie encore davantage les degrés de validité d'un message. Présentons sous forme d'une « table des jugements de vérité » les différentes combinaisons possibles entre un jugement initial catégorique (V = vrai ; F = faux) ou hypothétique (V? = incertitude quant à la vérité ; F? = incertitude quant à la fausseté) et un jugement final catégorique (voir la figure 1).

Figure 1. Table des jugements de vérité

Cas	Jugement initial	Jugement final
1	V	V
2	V	F
3	F	V
4	F	F
5	V?	V
6	V?	F
7	F?	V
8	F?	F

Commentons brièvement les huit cas présentés dans la figure 1. Tous les jugements initiaux portés sur un message sont explicites, sauf dans les jugements initiaux catégoriques de vérité, où ils sont laissés implicites en vertu du principe de coopération (cas 1 et 2). Les cas 2 et 4 sont des jugements initiaux erronés (ou mensongers, si l'émetteur est conscient de la fausseté de son message). Le cas 3 couvre tous les vrais récits de fiction, y compris les romans, où le jugement initial « F » est confirmé. Enfin, les cas 5 à 8 inclusivement se rapportent aux démarches de la vie quotidienne où les agents sont parfois obligés de fonder leur action sur des données incertaines, ainsi qu'à la démarche expérimentale fondée sur la production initiale d'une hypothèse et la confirmation ou l'infirmité ultérieure de cette hypothèse à la suite d'une épreuve.

Pour l'enseignement de la théorie du récit et de la sémiotique dans les ateliers de composition littéraire à l'école, ce sont les récits de fiction courts qui nous semblent les plus appropriés, et cela, pour deux raisons : la rapidité avec laquelle ils peuvent être produits et le plaisir qu'ils sont susceptibles de procurer à leur auteur, si les bases de données sont bien choisies.

1.3 Domaine d'application d'une base de données

Si les romans contiennent des données fictives, les romans en tant que tels sont des objets réels et leur titre peut figurer dans une base de données vraies. Les bibliothèques classent les romans sous diverses étiquettes, notamment par nom d'auteur, pays de provenance, époque et ce que les spécialistes en études littéraires appellent habituellement « sous-genre » du roman, par exemple « policier » (Aron, Saint-Jacques et Viala, 2002 : 526). Nous appellerons les sous-genres du roman des « genres d'histoires », ce qui permet de faire abstraction à la fois du support matériel des histoires racontées (voix, dessin, film, mime, image électronique, etc.) et du code (système des signes linguistiques ou iconiques). Le *Répertoire de vedettes-matière de l'Université Laval* renvoie à quelques dizaines de genres d'histoires au sens qui vient d'être donné. Citons en particulier les romans chrétiens, de guerre, d'espionnage, de science-fiction, écologiques, familiaux, féministes, judiciaires, médicaux, pastoraux, politiques et westerns. Ce répertoire est accessible sur abonnement (Hudon, 2013 : 105) ou sur demande particulière adressée à une bibliothèque universitaire du Québec. À chacun des genres d'histoires qui y sont mentionnés, et à chacun

de ceux qui peuvent exister dans toutes les littératures du monde, on peut associer une base de données fictives qui contiendrait les éléments plus ou moins fréquents des univers créés par l'imagination. Nous appellerons « domaine de la base de données fictives » l'ensemble des histoires racontées dans des œuvres de fiction appartenant à un genre d'histoires particulier, et qui font l'objet d'un dépouillement et d'un classement systématiques, et « catégories » les sous-ensembles de données ayant un caractère commun.

L'idée d'une base de données par genre d'histoires a été entrevue dès 1928 par Propp, auteur qui a inspiré à retardement plusieurs théoriciens français du récit des années 1960, notamment Greimas, mais pour une raison différente de celle qui nous occupe maintenant. Rappelons que l'étude de Propp porte sur un corpus d'une centaine de contes merveilleux russes : c'est le « domaine » des histoires merveilleuses russes produites à une époque lointaine. Souhaitant proposer un nouveau principe de classification des contes, Propp s'intéresse surtout aux catégories d'actions qui y apparaissent et qu'il appelle des « fonctions ». Pour chacune des fonctions, il donne toutefois des exemples où figurent des listes de personnages et d'objets bien précis. Voici quelques exemples regroupés sous la fonction « malveillance » : 1) enlèvement d'un être humain (fille du roi, fille d'un paysan, petit garçon, fiancée d'un frère plus jeune) par un agresseur (dragon, sorcière, frère aîné) ; 2) vol d'un objet magique (coffret, chemise, cheval) par un agresseur (« gars qui ne paye pas de mine », fille du roi, petit homme) (Propp, 1970 [1928] : 53-54). Si l'action de malveillance n'est pas propre au domaine des histoires merveilleuses, par contre, nous sentons que plusieurs des personnages ou objets mentionnés caractérisent ce genre d'histoires. Propp envisage en outre une analyse des attributs des personnages. Il suggère même quelques rubriques : dénomination, âge, sexe, situation, aspect extérieur, traits particuliers, formes particulières d'apparition, habitation (Propp, 1970 [1928] : 147-148). Même si Propp ne nous a pas fourni de listes de données pour les catégories d'attributs, les listes de personnages, d'actions et d'objets rangées sous les différentes fonctions du conte merveilleux constituent en somme une base de données qui pourrait servir, au moyen de recombinaisons judicieuses, à produire de nouvelles histoires merveilleuses.

1.4 Regroupement des données en catégories

Existe-t-il une liste finie et universelle de catégories pour les bases de données ? Comme premier exemple, prenons celui du répertoire de personnages de la littérature française établi par Ajame et Brucker (1981). À partir d'une analyse de plus de 200 romans de 83 écrivains, les auteurs du répertoire ont constitué une liste de quelques centaines de personnages qu'ils ont caractérisés au moyen de données, quand il y en avait, appartenant aux 28 catégories suivantes : nom, prénom, surnom, nationalité, époque, âge, domiciles, santé, aspect physique, habillement, famille, éducation et études, activités professionnelles, fortune, domesticité, voyages, vie sexuelle et sentimentale, amitiés, inimitiés, relations, opinions religieuses, opinions politiques, qualités, défauts, aime, n'aime pas, signes particuliers, mort.

Comme second exemple, prenons la liste des 16 catégories auxquelles se réfère explicitement Petitjean (1994) pour la production de robinsonnades, c'est-à-dire des histoires fictives de naufragés qui se déroulent dans un cadre exotique, comme l'histoire de Robinson Crusoé. Voici les catégories mentionnées çà et là par l'auteur dans son texte : abri, animaux, armes, bateaux, croyances, langage, mode de couchage, nourriture, paysage, physique des personnages, plantes, pratiques magiques, pratiques sexuelles, relations sociales, rites sacrificiels, travail. Les données appartenant à ces catégories n'étaient pas fournies. L'expérience pédagogique consistait, entre autres, à confier à des élèves de 5^e de collège (environ 13 ans) le soin de les extraire d'un petit corpus de robinsonnades et de recombinaison certaines de ces données pour créer de nouvelles histoires.

La comparaison de la liste des catégories d'Ajame et Brucker avec celle de Petitjean révèle quelques points d'une toute relative convergence : domiciles – abris ; activités professionnelles – travail ; relations – relations sociales ; vie sexuelle et sentimentale – pratiques sexuelles ; opinions religieuses – croyances. Tout comme les catégories d'Ajame et Brucker ne s'appliquent pas à tous les personnages de leur corpus par manque de données à leur sujet, certaines catégories peuvent être utiles à la production de certains genres d'histoires et d'autres non : une liste d'armes convient très bien s'il s'agit d'inventer une histoire policière, merveilleuse héroïque ou de science-fiction... ou une robinsonnade, mais très peu si la consigne d'écriture concerne une histoire familiale ou sentimentale.

1.5 Étapes de l'élaboration d'une base de données

Avant tout, il faut prendre en considération les destinataires de la base de données qui servira à la production d'histoires fictives. S'il s'agit de jeunes de moins de 18 ans qui participent à des ateliers de composition littéraire dans les écoles, le créateur de la base de données doit tenir compte des règles de la censure et restreindre les genres d'histoires et les données à mettre à la disposition des utilisateurs. Une fois le domaine ou le genre d'histoires choisi, le créateur de la base le nomme en utilisant la désignation la plus courante ou la plus claire et donne une définition de l'expression. L'analyse peut porter sur une dizaine d'œuvres fictionnelles que les experts du domaine jugent les plus représentatives. Elle peut inclure aussi des œuvres documentaires liées au genre traité. Par exemple, si l'analyste s'intéresse aux histoires maritimes, il pourrait tirer des données d'un livre sur les vrais pirates et flibustiers des Antilles au XVII^e siècle. Quant aux données extraites des œuvres, elles sont regroupées dans une vingtaine ou une trentaine de catégories. Dans chacune des catégories, le nombre de données pourrait osciller entre quelques-unes et plusieurs dizaines. Les données et les catégories sont présentées en ordre alphabétique.

Prenons l'exemple de la base de données « Histoires exotiques » (Noël-Gaudreault et Gaudreault, 2013 : 33-40). Les créateurs de la base définissent d'abord d'une façon précise ce genre d'histoires : « Nous entendons par "histoire exotique" une histoire fictive qui se passe à notre époque, hors de l'Amérique du Nord et de l'Europe, mais dont l'auteur et au moins le personnage principal résident dans un pays nord-américain ou européen » (Noël-

Gaudreault et Gaudreault, 2013 : 29). Ensuite, ils présentent 19 catégories, parfois subdivisées, dans lesquelles sont répartis 749 éléments d'information. Voici des extraits de listes d'éléments appartenant à deux catégories :

BUTS. Chercher la carcasse d'un avion dans la jungle. Échapper à des hommes déguisés qui attaquent. Participer à une cérémonie vaudou pour jeter un sort à quelqu'un. Partir à la recherche d'un trésor. Retrouver un camion de dynamite qui a disparu. Sauver des amis retenus en otages. Se libérer d'un piège. S'évader d'une prison. Survivre après une panne d'avion en pleine jungle.

NATURE. Arbre à pain, bambou, buisson épineux, cactus, cataractes, désert de sable, eau couleur de jade, forêt marécageuse, hévéa, île frangée de cocotiers, monolithes de basalte rouge.

Les listes partielles de données ci-dessus ont été extraites de romans d'Henri Vernes, auteur belge qui a produit une série d'aventures exotiques de son héros Bob Morane.

2. Utilisation d'une base de données pour la production de récits de fiction

Dans cette deuxième partie, nous établirons d'abord la liste des opérations de production d'un message, notamment d'un récit de fiction, et la comparerons à celle des opérations de production de messages verbaux de Jakobson et à celle des opérations de calcul automatique de Turing, qui aboutissent à la production d'un message dont le contenu est un nombre. Ensuite, nous décrirons les différents degrés de valeur d'une opération et comparerons cette conception de la valeur à celle prônée par Greimas. Enfin, nous indiquerons dans quelle mesure les notions de quantité et d'ordre associées à la définition des opérations de production de messages correspondent aux notions fondamentales de la théorie du récit de Genette.

2.1 Opérations

Considérons les mouvements possibles d'un objet entre deux lieux contigus que nous nommerons « A » et « B ». L'existence de ces deux lieux implique deux sortes de mouvements : un mouvement de A vers B et un autre de B vers A. Appelons « opération de retrait » le mouvement d'un objet de A vers B provoqué par un agent et « opération d'introduction » le mouvement de B vers A provoqué par un agent. De tels mouvements contraires peuvent ne pas se produire, l'objet étant maintenu en repos par l'agent dans le premier lieu ou n'étant pas introduit dans le second lieu après avoir été extrait du premier.

Nous appellerons ces deux mouvements contraires qui ne se réalisent pas « opération de non-retrait » et « opération de non-introduction »¹.

L'agent qui produit un message à l'aide d'une base de données exécute successivement des opérations du genre mentionné ci-dessus : retraits de la base de données suivis d'introductions dans le message. Par suite des introductions successives, la quantité de données augmente dans le message. Elle ne diminue pas pour autant dans la base, car l'agent effectue au préalable une copie des données. Cependant, dans le message, la quantité de données peut diminuer par suite de retraits hors du message. Par ailleurs, les retraits et les introductions peuvent être effectués à l'intérieur même du message sans que changent ni la quantité ni la nature des données. Dans ce cas, les opérations ont une incidence éventuelle seulement sur l'ordre des données.

Le mécanisme de production de messages au moyen des opérations élémentaires proposées ci-dessus est connu depuis longtemps, même si les termes utilisés par les auteurs ne sont pas les mêmes. Pour le linguiste Jakobson (1963 [1956] : 48), le locuteur produit un message en sélectionnant des mots [retrait] dans un répertoire lexical et en les combinant [introduction] dans des phrases. Parfois, sans changer le sens de son message, le locuteur remplace un signe linguistique par un signe équivalent, mais plus clair pour le destinataire. Selon Jakobson (1963 [1959] : 79-80), il existe trois types de remplacement de signes linguistiques équivalents : 1) la traduction intralinguale, ou remplacement de mots par des mots équivalents de la même langue (synonymes) ; 2) la traduction interlinguale, ou remplacement de mots par des mots équivalents d'une autre langue ; 3) la traduction intersémiotique, ou remplacement de mots par des signes non linguistiques équivalents, ou l'inverse. Dans tous les cas de traduction, il s'agit de retirer un signe d'un message, de retirer la copie d'un signe équivalent d'un répertoire et d'introduire ce signe équivalent dans le message. La description du processus de traduction présentée par Jakobson n'exige pas le recours à d'autres opérations que celles de retrait et d'introduction. Elle fait toutefois ressortir la distinction entre le support ou la forme d'une donnée et son contenu ou son sens, qui est au fondement de la notion de signe équivalent : un signe qui n'a pas la même forme qu'un autre signe, mais qui a le même sens.

Les opérations de retrait et d'introduction proposées pour expliquer le processus de production de messages en général se retrouvent aussi sous un autre nom dans un travail du mathématicien Turing, antérieur d'une vingtaine d'années à celui de Jakobson, mais apparemment sans influence sur ce dernier. Pour résoudre un problème propre à la mathématique de l'époque, Turing conçoit sur papier un calculateur automatique, qu'on appellera plus tard la « machine de Turing ». La machine comprend trois parties : 1) une bande divisée en cases où sont inscrits des symboles ; 2) un dispositif mobile de lecture-écriture ; 3) un ensemble d'instructions. Le dispositif mobile inspecte les cases et, selon ce qu'il y trouve et les instructions qui s'appliquent, exécute différentes opérations.

¹ La conception de ces quatre opérations élémentaires doit beaucoup à l'étude des relations contraires et contradictoires de Greimas. Elle s'en démarque toutefois très nettement (voir Gaudreault, 2008).

Voici les quatre opérations élémentaires du dispositif mobile : 1) déplacement d'une case vers la droite ; 2) déplacement d'une case vers la gauche ; 3) effacement d'un symbole dans la case inspectée ; 4) inscription d'un symbole dans la case inspectée (Turing, 1995 [1936] : 54). Sous l'opération d'effacement d'un symbole, nous reconnaissons l'opération de retrait et, sous l'opération d'inscription d'un symbole, celle d'introduction. Quant aux déplacements vers la gauche ou la droite, ils signalent l'importance de l'ordre des éléments à l'intérieur du message. Turing ajoute une dernière opération à ses quatre opérations de calcul automatique : la recopie d'une séquence de symboles à la fin de la bande (Turing, 1995 [1936] : 61). Cette opération accessoire est apparentée à une autre proposée plus haut : la copie d'une donnée inscrite dans une base de données.

2.2 Valeurs

Un jugement de valeur s'ajoute au jugement de réalité dont nous avons parlé au début à propos de la validité des données. L'importance ou la valeur d'un message, et par conséquent des données qu'il contient, varie selon le lien que celui-ci entretient avec les objets qui satisfont les besoins du producteur du message. Deux besoins permanents orientent l'action d'un être vivant : 1) un besoin de conservation l'incite à rechercher tout objet utile à son fonctionnement ; 2) un besoin de protection le pousse à éviter ou à éliminer tout objet nuisible à son fonctionnement². Outre des objets utiles ou nuisibles, il peut en trouver d'inutiles ou de non nuisibles. Par conséquent, nous dirons qu'un objet a quatre valeurs possibles pour un être vivant : 1) une valeur positive, s'il lui est utile ; 2) une valeur négative, s'il lui est nuisible ; 3) une valeur non positive, s'il ne lui est pas utile ; 4) une valeur non négative, s'il ne lui est pas nuisible.

Par « objet », il faut entendre un être non vivant, y compris un message, un être vivant ou une association d'êtres vivants. Une association d'êtres vivants comporte les mêmes organes qu'un être vivant individuel : des organes de commande, des organes d'exécution de la commande et des organes d'information sur l'exécution de la commande (Wiener, 1971 [1950 et 1954] : 88-89 ; 164-165). Prenons l'exemple d'un État : le gouvernement correspond au cerveau de l'être vivant ; l'armée, la police, les sociétés agricoles, industrielles, commerciales et autres, qui satisfont les besoins de protection et de conservation, correspondent aux muscles et à des organes physiques spéciaux ; les organismes de surveillance comme les tribunaux correspondent aux organes sensoriels. En conséquence, tout le fonctionnement des êtres vivants et des associations d'êtres vivants repose sur la production et la communication de deux types de messages : les messages de commande et les messages d'information³. Dans le cas d'un État, les lois sont des messages qui commandent d'exécuter ou de ne pas exécuter certaines actions (prescriptions et interdictions). Parmi les interdictions, on compte les limites imposées au contenu des messages qui peuvent être échangés entre des agents. Comme le non-respect des lois entraîne

² Résultat de l'étude de la théorie des besoins de Maslow (dans Gaudreault, 2007).

³ Les récits sont des messages informatifs.

des sanctions, c'est-à-dire des actions qui ont une valeur négative pour le contrevenant, un agent évitera d'introduire dans ses messages des éléments de contenu (des données) susceptibles d'entraîner une sanction et qui, de ce fait, prennent une valeur négative pour lui. Par exemple, pour la presse et le cinéma, il existe des interdictions touchant l'expression de la sexualité (bonnes mœurs, érotisme, obscénité) et l'agressivité (ordre public, violence, incitation au crime) (Michard, 1967 : 55).

La non-introduction d'un élément de valeur négative dans un message, dont nous venons juste de parler à propos du respect de certaines interdictions, n'est que l'une des 16 combinaisons possibles entre les quatre opérations de production de message et les quatre valeurs attribuables par un agent à des objets quelconques, y compris des messages et des données. La figure 2 regroupe l'ensemble de ces combinaisons (voir ci-dessous).

Figure 2. Combinaison des opérations et des valeurs

	P	N	nP	nN
R	RP	RN	RnP	RnN
I	IP	IN	InP	InN
nR	nRP	nRN	nRnP	nRnN
nI	nIP	nIN	nInP	nInN

Code : R = retrait ; I = introduction ; nR = non-retrait ; nI = non-introduction ; P = objet de valeur positive ; N = objet de valeur négative ; nP = objet de valeur non positive ; nN = objet de valeur non négative

Parmi les 16 combinaisons d'opérations et de valeurs, certaines ont déjà été décrites par des théoriciens du récit, en particulier Greimas et Genette. Greimas établit quatre types de transferts d'objets entre sujets⁴ : le don, le contre-don, l'échange et l'épreuve (Greimas, 1983 [1973] : 39-40). Dans tous les cas, il s'agit de la même sorte d'objets : des êtres non vivants, mais non des messages, qui ont une valeur positive pour les sujets (un trésor, de la nourriture). Le don est le retrait volontaire d'un objet de valeur positive des possessions d'un premier sujet suivi de l'introduction de cet objet parmi les possessions d'un second sujet (RP suivi d'IP). Le contre-don consiste dans la même suite d'opérations, concernant le même objet, effectuées par le second sujet au profit du premier. L'échange est la suite d'opérations du don du premier sujet répétées par le second sujet, mais concernant un objet différent de l'objet donné par le premier sujet. L'épreuve est ce que nous pourrions appeler un « don forcé », c'est-à-dire le retrait forcé d'un objet de valeur positive des possessions d'un premier sujet et l'introduction de cet objet parmi les possessions d'un second sujet (voleur ou agresseur). Les quatre types de transferts n'illustrent que 2 des 16 combinaisons (RP et IP). Toutefois, Greimas apporte deux couples de distinctions nouvelles : 1) même objet et objet différent ; 2) opération volontaire et opération forcée.

⁴ Greimas désigne aussi ces transferts comme des disjonctions et des conjonctions de sujets et d'objets, c'est-à-dire des opérations analogues à celles de retrait et d'introduction d'éléments effectuées dans une base de données ou dans un message.

Quant à Genette, autre grand théoricien du récit et contemporain de Greimas, il conçoit ce que nous appelons une opération de non-introduction de données dans le message sous le nom d'« ellipse », c'est-à-dire « omission » (Genette, 1972 : 92). En ce qui a trait aux données introduites, l'auteur les classe d'abord en pensées, perceptions, paroles et actions des personnages et distingue diverses « focalisations » (Genette, 1972 : 206-211), selon que certains types de données sont présents ou absents (présence de pensées et de perceptions, ou focalisation interne ; absence de pensées et de perceptions, ou focalisation externe) ou selon que l'auteur introduit ou non une plus grande quantité de données à propos de certains personnages par rapport aux autres (focalisation fixe ou variable, absence de focalisation). Ensuite, d'autres données concernent des objets matériels ou le décor des événements, qui font l'objet de « pauses descriptives » (Genette, 1972 : 129, 133). Genette souligne en outre l'importance de l'ordre d'exposition des événements et reprend les distinctions traditionnelles entre les rétropections, ou retours en arrière, et les anticipations (Genette, 1972 : 82). Selon notre approche, ces transformations de l'ordre d'exposition des événements sont des opérations successives de retrait et d'introduction de données à un autre endroit à l'intérieur du message.

Conclusion

La création de bases de données pour la production de récits de fiction dans les ateliers de composition littéraire faciliterait l'enseignement de la sémiotique conçue comme la science des signes et de leur emploi dans la production et la communication de messages, y compris des récits. Les bases de données sont des répertoires auxquels on accède soit à l'aide d'un ordinateur, soit manuellement. Chaque répertoire a un domaine d'application, par exemple, les genres d'histoires fictives que véhiculent les œuvres littéraires d'un pays à une certaine époque. Les données du répertoire sont classées en catégories dont le nombre et la nature varient selon le genre d'histoires.

Pour produire un récit de fiction à l'aide d'une base de données, seulement quatre opérations élémentaires sont requises : le retrait (après une opération accessoire de copie) ou le non-retrait d'une donnée de la base et l'introduction ou la non-introduction de la donnée dans le message. Les données ont plusieurs valeurs possibles : elles peuvent être utiles, nuisibles, non utiles ou non nuisibles au but poursuivi. Le producteur d'un message doit éviter d'introduire dans celui-ci des données qui lui sont nuisibles, notamment celles qui font l'objet d'interdictions de la part des autorités des différents groupes sociaux auxquels il appartient (État, famille, école, entreprise et autres associations). Quant aux données utiles, le producteur d'un message continue d'en introduire jusqu'à ce qu'il ait atteint la quantité voulue pour chacun des objets dont il veut parler. Cela fait, il examine le contenu du message et maintient ou modifie l'ordre des données introduites. S'il décide de modifier cet ordre, il a de nouveau recours aux opérations élémentaires qui lui ont déjà servi à produire le message à partir de la base de données.

Le recours à des bases de données pour faciliter la production de récits de fiction dans les ateliers de composition littéraire à l'école a deux conséquences pratiques importantes : premièrement, la fin de la prééminence accordée à la mémoire interne du scripteur pour la production de récits de fiction, car la base de données joue le rôle de mémoire auxiliaire externe ; deuxièmement, la création d'un nouveau débouché et le recours à de nouvelles méthodes pour les spécialistes en études littéraires qui sont aussi des sémioticiens. Dans le présent travail, nous n'avons qu'effleuré l'aspect quantitatif des bases de données quand nous avons dit que celles-ci regroupent des éléments « plus ou moins fréquents » des univers créés par l'imagination. Pour approfondir la question de la fréquence des éléments rapportés sous forme de données dans une base, il conviendrait d'étudier attentivement l'utilisation de la théorie des probabilités, notamment les probabilités de transition, dites « chaînes de Markov », par Shannon dans sa théorie mathématique de la communication, œuvre que Moles (1975 : 11) a qualifiée d'« événement intellectuel ».

Notice biobibliographique

Romain Gaudreault est chercheur indépendant. En 1991, il a obtenu un diplôme de doctorat de l'Université Laval pour une thèse intitulée « Jalons pour une sémiotique fondamentale ». Il a publié notamment « Renouveau du modèle actantiel » (revue *Poétique*, 1996), « Valeurs, émotions, sentiments et crise » (revue *Semiotica*, 1998) et « Reflets des tensions nationales dans l'art public de Montréal » (dans Hébert et Guillemette, *Performances et objets culturels*, 2010). Il a également codirigé, avec Daniela Roventa-Frumusani, *Pour connaître la science des signes* (Craiova, Roumanie, 2001).

Ouvrages cités

- AJAME, P. et M. BRUCKER (1981), *300 héros et personnages du roman français. D'Atala à Zazie*, Paris, Balland.
- ARON, P., D. SAINT-JACQUES et A. VIALA (dir.) (2002), *Le dictionnaire du littéraire*, Paris, Presses universitaires de France.
- CALVINO, I. (1984 [1967]), « Cybernétique et fantasmes, ou de la littérature comme combinatoire », dans *La machine littérature. Essais*, Paris, Seuil p. 11-29.
- GAUDREAU, R. (2007), « Source du plaisir et sémiotique », dans L. HÉBERT (dir.), *Le plaisir des sens. Euphories et dysphories des signes*, Québec, Presses de l'Université Laval (coll. « Vie des signes »), p. 69-81.
- GAUDREAU, R. (2008), « Théorie du récit et sémiotique : apport d'A. J. Greimas et nouvelles propositions », *Semiotica*, n° 172 (1/4), p. 1-9.
- GENETTE, G. (1972), *Figures III*, Paris, Seuil.
- GREIMAS, A. J. (1983 [1973]), « Un problème de sémiotique narrative : les objets de valeur », dans *Du sens II. Essais sémiotiques*, Paris, Seuil p. 19-48.

- GRICE, P. (1989), *Studies in the Way of Words*, Cambridge, Massachusetts/Londres, Harvard University Press.
- HUDON, M. (2013), *Analyse et représentation documentaires. Introduction à l'indexation, à la classification et à la condensation des documents*, Québec, Presses de l'Université du Québec.
- JAKOBSON, R. (1963), « Deux aspects du langage et deux types d'aphasie », dans *Essais de linguistique générale*, traduit de l'anglais et préfacé par Nicolas Ruwet, Paris, Éditions de Minuit 1956, p. 43-67 .
- JAKOBSON, R. (1963), « Aspects linguistiques de la traduction », dans *Essais de linguistique générale*, traduit de l'anglais et préfacé par Nicolas Ruwet, Paris, Éditions de Minuit 1956, p. 78-86.
- LEJEUNE, P. (1975), *Le pacte autobiographique*, Paris, Seuil.
- MICHARD, H. (1967), « Contrôle cinématographique et protection des jeunes », *Communications*, n° 9, p. 54-63.
- MOLES, A. (1972 [1958]), *Théorie de l'information et perception esthétique*, Paris, Denoël.
- MOLES, A. (1975), « Préface », dans C. E. SHANNON et W. WEAVER, *Théorie mathématique de la communication*, traduit de l'anglais [*The Mathematical Theory of Communication*, 1949] par J. Cosnier, G. Dahan et S. Economidès, Paris, Retz – C.E.P.L., p. 11-27.
- MORVAN, P. (dir.) (1996), *Dictionnaire de l'informatique*, Paris, Larousse.
- NOËL-GAUDREAULT, M. et R. GAUDREAULT (2013), *Outils pour écrire des récits d'aventures de science-fiction et d'aventures exotiques*, Montréal, Éditions du Bégonia voyageur.
- PETITJEAN, A. (1994), « Écrire des récits d'aventures », *Pratiques*, n° 83 (septembre), p. 79-124.
- PROPP, V. I. (1970), *Morphologie du conte*, traduit du russe [*Morfologija skazki*, 1928 et 1966] par Claude Ligny, Paris, Gallimard.
- TURING, A. (1995 [1936]), « Théorie des nombres calculables, suivi d'une application au problème de la décision », article traduit de l'anglais et annoté par Julien Basch, dans A. TURING et J.-Y. GIRARD, *La machine de Turing*, Paris, Seuil, p. 47-104.
- WIENER, N. (1971), *Cybernétique et société. L'usage humain des êtres humains*, édition synoptique fondée sur la traduction française des deux éditions anglaises de 1950 et de 1954, Paris, 10/18.