

Les métaphores dans les sciences de la vie et le situé socioculturel

La compréhension de nouveaux domaines de recherche et les connaissances scientifiques se développent notamment par le biais du raisonnement métaphorique. L'analogie avec des domaines d'études déjà connus permet d'inventer une nouvelle compréhension scientifique ou d'en faciliter la compréhension dans un but de vulgarisation. L'étude des lexicalisations métaphoriques relevées par l'auteur dans un corpus de textes anglais relatifs aux sciences de la vie montre la projection du cadre conceptuel de l'information qui est à l'œuvre dans l'étude de l'ADN : le matériel peut être conçu tantôt comme un langage, un atlas, un programme informatique ou un film. Ce modèle métaphorique est situé sur le plan socioculturel et mérite l'attention des terminologues et des traducteurs qui sont inévitablement confrontés à la question de l'universalité de la métaphore, et, partant, de la traduction des lexèmes métaphoriques.

Termes-clés :

néologisme; analogie; lexicalisation métaphorique, terminologie sociocognitive; terminologie plurilingue.

1 Introduction

Dans le cadre de l'étude des néologismes d'un corpus de textes en anglais portant sur les sciences de la vie, nous avons été frappée par le nombre élevé des lexicalisations métaphoriques. Notre recherche d'une explication nous a amenée à cerner de plus près le modèle métaphorique qui semble sous-tendre la nouvelle compréhension des principes de la vie ainsi que le développement des sciences de la vie. En première partie de cet article, nous présentons au lecteur quelques exemples de néologismes métaphoriques tirés du corpus étudié. Dans sa deuxième partie, nous présentons le modèle métaphorique multiniveau qui semble servir de base à la compréhension de la vie en tant qu'information. Les différents types de textes étudiés nous ont montré la nécessité de différencier le potentiel créatif du potentiel didactique du modèle de l'information, ainsi que nous l'expliquons dans une troisième partie. Dans le quatrième volet de notre article, nous proposons de sensibiliser traducteurs et terminographes à l'apport du modèle métaphorique dans le cadre de leur formation. Nous y affirmons aussi notre conviction que la théorie de la terminologie pourrait

tirer avantage d'études de cas approfondies pour élaborer des méthodologies plus variées et situées lors de la mise en place de différents types de bases de données terminologiques. La méthodologie élaborée en terminologie sociocognitive pour l'étude de la conceptualisation dans les langues de spécialité en plusieurs langues pourrait, en parallèle, fournir des études de cas intéressantes pour les sciences cognitives. Enfin, dans la cinquième et dernière partie de l'article, nous pointons le fait que la cognition est nécessairement enchâssée (*embedded*) dans un situé social et culturel, dimension qui a bénéficié récemment d'un intérêt marqué des sciences de la cognition. La cognition est définie comme un système culturellement réparti qui émerge de l'interaction entre les membres d'un groupe culturel. Dans notre cas, les groupes culturels concernés seraient ceux formés par les chercheurs dans le domaine de la technologie de l'ADN recombinant.

2 Néolexicalisations métaphoriques dans les sciences de la vie

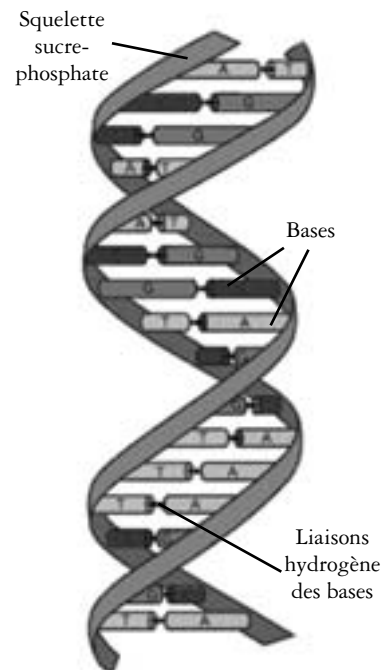


Figure 1. Structure en double hélice de l'ADN

De nos jours, l'idée que l'ADN est le support de l'information génétique est si ancrée dans la pensée biologique qu'il nous est difficile d'imaginer l'ampleur du saut conceptuel qu'elle a représenté à l'époque où les fondements en ont été découverts. Ce n'est qu'en 1953 que Watson et Crick proposèrent un modèle descriptif de la structure de l'ADN fondé sur une analogie. Ils découvrirent que l'ADN, que l'on savait constitué de quatre bases – l'adénine (A), qui s'associe toujours à la thymine (T), et la guanine (G), qui s'associe toujours à la cytosine (C) – avait une structure en forme de double hélice (fig. 1). L'ordre des paires de bases (ou des nucléotides) fournit l'information nécessaire pour assurer la réplication de l'ADN et pour rendre la traduction de l'ADN en protéine possible.

Watson, Crick et leurs collègues chercheurs ont commencé à comprendre l'ADN en termes d'information et ce modèle analogique a donné naissance à toute une série de néologismes métaphoriques. Dans les extraits suivants, nous pouvons les observer (en lettres grasses) *in vivo* :

*How can a long chain of nucleotides **encode** the instructions for an organism or even a cell? And how can these **messages be copied** from one generation of cells to the next? The answers lie in the structure of the DNA molecule (Alberts et al., 1994:99)*

*It therefore seems that the basic **unit of information** in the **genetic code** can be no smaller than a triplet of bases. But there are 64 triplets--more than three times the number needed. Explaining away this excess became a major preoccupation of coding theorists. (Hayes, 1998)*

*1)The **genetic code** is read in a sequential manner starting near the 5' end of the mRNA. This means that **translation** proceeds along the mRNA in the 5' --> 3' direction which corresponds to the N-terminal to C-terminal direction of the amino acid sequences within proteins.*

*2)The **code** is composed of a triplet of nucleotides.*

*3)That all 64 possible combinations of the 4 nucleotides code for amino acids, i.e. the **code** is degenerate since there are only 20 amino acids. The precise **dictionary** of the genetic code was determined with the use of **in vitro translation systems** and polyribonucleotides. The results of these experiments confirmed that some amino acids are **encoded** by more than one triplet codon, hence the degeneracy of the **genetic code**. These experiments also established the identity of **translational termination codons**. ([http://web.indstate.edu/thcme/mwking/protein-synthesis.html#genetic code](http://web.indstate.edu/thcme/mwking/protein-synthesis.html#genetic%20code))*

*DNA is not **translated** directly to protein. Instead, DNA is transcribed into mRNA, which is then **edited and translated** into*

*proteins by ribosomal RNA (rRNA). The RNA is **transcribed** from DNA (**transcription**) in the nucleus, and then exits into the cytoplasm, where rRNA **translates** the mRNA into polypeptide chains (**translation**), which then form proteins. These two main steps are called transcription and translation; transcription occurring in the nucleus and translation in the cytoplasm. (<http://sidwell.edu/us/science/vlb98/projects/protein/smarks/protein/index.html>)*

À l'exception du terme « codon » (cf. partie 3.1), tous les néologismes présents dans ces extraits résultent d'un transfert métaphorique du **domaine de l'information** : *to encode* (coder), *message, to copy* (copier), *unit of information* (unité d'information), *genetic code* (code génétique), *dictionary* (dictionnaire), *translation system* (système de traduction), *to translate* (traduire), *to transcribe* (transcrire), *to edit* (éditer). Examinons à présent de plus près le modèle métaphorique mis en œuvre.

3 Le modèle métaphorique

Dans cette partie, nous nous concentrerons sur le développement de la compréhension et des connaissances par le biais du raisonnement métaphorique (la projection d'un cadre conceptuel source sur un cadre conceptuel cible). Le raisonnement métaphorique conduit à la compréhension d'un fait, d'une situation, d'un procès ou de tout type de catégorie nouveau au départ de l'analogie imaginée entre ce que l'on tente d'appréhender, de comprendre, et quelque chose que l'on connaît et comprend déjà. Cette faculté d'invention ou de création prend corps et laisse ses traces dans la lexicalisation de néologismes. Le raisonnement imaginaire, dont le produit est la dénomination métaphorique de nouvelles catégories au moyen de lexèmes existants, s'enracine dans l'expérience humaine.

Les disciplines des sciences de la vie se développent dans des cadres cognitifs qui peuvent être vus comme des *gestalts* (Lakoff et Johnson, 1980 ; Lakoff, 1987). Pour élaborer de nouvelles idées, les praticiens du domaine des sciences de la vie accaparent des cadres cognitifs existants dans des domaines d'expérience étrangers aux sciences de la vie (les domaines sources). Les modèles métaphoriques laissent leurs traces dans les lexicalisations de néologismes.

Le modèle métaphorique fonctionne comme une gestalt. Nous pouvons l'observer si nous comparons la part de la gestalt métaphorique qui débouche sur une lexicalisation à la part susceptible d'être exprimée sous la forme de propositions lexicalisables quand davantage d'éclaircissements sont requis, c.-à-d. dans des circonstances où des explications supplémentaires sont nécessaires, par ex. dans le cadre de la formation de futurs spécialistes (cf. partie 5).

Une série d'analogies qui ont laissé leur trace dans des lexicalisations métaphoriques fournissent le soubassement des progrès de la compréhension des mécanismes de la vie. Il est possible de constater le développement d'un nouveau champ lexical reposant sur les structures d'une gestalt fondée sur un rapport métaphorique ou analogique spécifique.

Le processus fait appel à plusieurs sous-domaines du modèle de l'information, lesquels permettent d'expliquer nombre de lexicalisations métaphoriques (fig. 2), dont :

- L'ADN est un langage. Les gènes sont des messages écrits dans une langue. Il s'agit d'un premier sous-domaine qui s'appuie sur l'expérience du fait que l'information s'exprime souvent par le biais du langage (section 3.1).
- L'ADN est un atlas. La totalité de l'information génétique d'un organisme (le génome) peut être cartographiée. Les généticiens indiquent l'emplacement des gènes sur des cartes génétiques de la même manière que les explorateurs de notre planète indiquaient sur des cartes l'information relative à l'emplacement de faits géographiques qu'ils avaient observés (section 3.2).
- L'ADN est un programme informatique exécutable par la cellule. Ce troisième sous-domaine repose sur l'expérience du fait que l'information est souvent stockée et disponible sous une forme électronique (section 3.3).
- L'ADN est un film susceptible d'être « lu » par un projecteur. Ce quatrième sous-domaine s'appuie sur l'expérience du fait que l'information peut être conservée et mise à disposition sur de la pellicule cinématographique (section 3.4).

| Domaine source de l'analogie | Sous-domaines du modèle de l'information |
|------------------------------|--|
| Modèle de l'information | Le matériel génétique (ADN) peut être conçu comme un langage (section 3.1). |
| | La totalité du matériel génétique d'un organisme (le génome) peut être conçue comme un atlas géographique (section 3.2). |
| | Le matériel génétique (ADN) peut être conçu comme un programme informatique (la cellule) (section 3.3). |
| | Le matériel génétique (ADN) peut être conçu comme de la pellicule cinématographique (section 3.4). |

Figure 2. Représentation des sous-domaines du domaine source (l'information) utilisé pour la compréhension de la génétique moléculaire

3.1 L'ADN est un langage

Un premier sous-modèle repose sur l'expérience du fait que l'information emprunte souvent la forme du langage. Dans ce cas, le matériel génétique (l'ADN) peut être compris comme s'il équivalait à un langage. Shapiro (1991 :4) décrit comment l'idée que notre plan d'ensemble est stocké sous la forme de lettres s'est imposée depuis un peu plus d'une

génération. Auparavant, nos ancêtres étaient inspirés par d'autres analogies imaginées qui ont laissé leur trace dans la langue. Les reliquats d'une idée bien plus ancienne marquent aujourd'hui encore de leur empreinte la langue comme des sortes de fossiles verbaux, à savoir l'idée que l'hérédité est conservée et transmise par le sang. L'idée est devenue si familière que nous ne réagissons pas à la lecture d'expressions telles que *sang royal*, *pur-sang*, *liens du sang*, *sang bleu* ou encore

sang mêlé. On doit la théorie liant hérédité et sang à Aristote (384-322 av. J.-C.) et à d'autres penseurs de cette époque. Pour les tenants de cette idée, l'hérédité impliquait un mélange des qualités parentales semblable au mélange de deux liquides différents.

Il n'en va pas de même pour du texte linéaire. Celui-ci peut être assemblé par « collage » (ou « épissage », *splicing*) mais non mélangé au sens propre. Un enfant reçoit une sélection de composants de l'un et l'autre parent. Certains traits sont préservés, d'autres traits d'un des parents peuvent être totalement perdus.

Nous résumons brièvement ci-dessous l'histoire de la théorie moderne de l'hérédité et indiquons en gras les lexicalisations fondées sur l'analogie sous-jacente avec le langage (notre résumé s'inspire des ouvrages d'Hodson (1992) et de Berg et Singer (1992)).

Les mécanismes de l'hérédité ont été découverts en 1866 par Gregor Mendel. Pour Mendel, les « facteurs » de l'hérédité (que nous appelons aujourd'hui « gènes ») étaient des entités purement abstraites. À l'époque, le facteur de grossissement des microscopes était de 1000 x ou plus. On identifia dans le noyau cellulaire des structures ressemblant à des bâtonnets colorés que l'on nomma « chromosomes » (du grec *khrôma* (coloré) et *sôma* (corps)). On comprit immédiatement qu'ils représentaient concrètement les « facteurs » abstraits de Mendel. Les chromosomes devinrent le point focal de la recherche en génétique. S'il était évident que les chromosomes fournissaient le support physique des mécanismes de l'hérédité imaginés par Mendel, ils ne pouvaient être assimilés aux facteurs mendéliens (gènes) pour la bonne raison qu'ils n'étaient pas assez nombreux.

Quoique le comportement des gènes fût de mieux en mieux compris, on ne disposait pas d'information sur leur comportement physique hors le fait qu'ils étaient ordonnés linéairement le long des chromosomes. Il était clair que les gènes étaient porteurs d'une manière ou l'autre de **messages** et que pour se perpétuer au-delà de la division cellulaire, ils devaient être capables de se **dupliquer** fidèlement. Toutefois, aucune théorie n'expliquait de quelle manière la nature des chromosomes leur conférait ces propriétés remarquables. Des expériences ultérieures identifièrent l'ADN (acide désoxyribonucléique) comme la clé du mystère. Francis Crick et James Watson résolurent, eux, le problème de la double structure hélicoïdale de l'ADN en construisant un modèle

réduit en fil de fer et carton. La molécule se présentait comme un escalier en colimaçon dont les marches sont faites de paires de bases (A (adénine) toujours associé à T (thymine), C (cytosine) toujours associé à G (guanine)) et dont les rampes sont des chaînes sucre-phosphate.

La découverte de Crick et Watson est que l'ADN a une structure qui lui permet de **se copier lui-même**. Vu que A doit être apparié avec T et C avec G, il s'ensuit que si une molécule d'ADN est découpée par le milieu dans le sens de la longueur, la totalité de l'information nécessaire pour reconstituer la molécule complète est présente.

La structure de l'ADN proposée par Crick et Watson fournit également une explication sur la manière dont un gène fonctionne du point de vue chimique. On savait déjà qu'un gène contrôle la production d'une protéine. Le modèle Crick-Watson montre comment le gène peut contenir un **message codé dans sa séquence de lettres**, message qui fournit les spécifications de la protéine à synthétiser. La structure codante de l'ADN remplit deux fonctions :

- permettre la **réplication** de l'ADN pour produire davantage d'ADN ;
- permettre la **transcription** de l'ADN en ARN, lequel est ensuite **traduit** en une protéine.

Transcription et traduction

Quand l'ADN produit une protéine, il le fait via une molécule intermédiaire très similaire à l'ADN, l'ARN (acide ribonucléique). L'ARN se présente en général uniquement sous la forme d'un brin simple. Une molécule d'ARN peut être synthétisée de manière répétée comme une copie conforme de la séquence de bases d'un brin d'ADN par une opération de **lecture** de la séquence concernée ; c'est ce processus qui est nommé **transcription**. Le brin d'ARN ainsi synthétisé est dénommé **ARN messager** (ARNm) car il est porteur du message indiquant qu'une protéine donnée doit être produite à partir du brin d'ADN qui a été copié.

Sydney Brenner travailla avec Crick au **déchiffrement du code ADN** et ils découvrirent qu'il est écrit en « mots » de **trois lettres**. Étant donné que les bases, au nombre de quatre, sont groupées par trois, on dispose de 64 (soit 4³) combinaisons possibles pour coder un nombre d'acides aminés limité, lui, à vingt. Il s'est avéré que la plupart des acides aminés peuvent être codés par plusieurs **codons** et que trois codons ne codent aucun acide aminé mais sont des signaux

d'arrêt (« stop ») qui indiquent la fin du **message codant la protéine** (fig. 3).

La **séquence de code** du gène ADN est utilisée pour construire une molécule d'ARN messager ; celle-ci est assemblée par une enzyme appelée ARN polymérase et c'est cette partie du processus que l'on nomme **transcription**.

Il existe par ailleurs de nombreuses molécules d'ARN d'un type différent, dénommées ARN de transfert (RNA_t). Les

molécules ARN_t sont uniquement constituées de trois bases. Ces trois bases forment un anticodon et chacune s'apparie avec la base du codon complémentaire de l'ARN_m. Cette partie du processus, dénommée **traduction**, intervient dans les ribosomes, lesquels agissent comme une sorte de cadre d'assemblage pour la construction des protéines. Les molécules d'ARN_t forment une ligne et les acides aminés se lient dans l'ordre spécifié pour former la chaîne protéique.

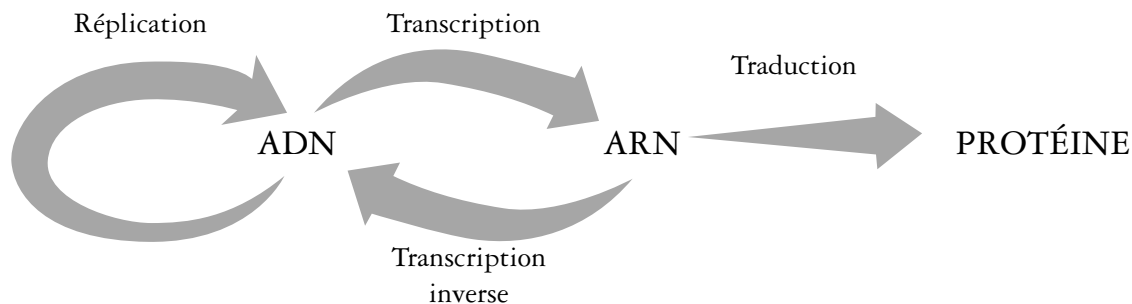


Figure 3. Les flèches montrent les processus et le sens de transfert de l'information génétique de l'ADN à l'ARN, de l'ARN à la protéine et de l'ARN à l'ADN (Berg et Singer, 1992:35).

Il est difficile de déterminer si le cadre cognitif ou la gestalt du « langage » est activé après que l'analogie entre gènes et langage est comprise ou si un détail de l'analogie – la visualisation des bases sur les brins d'ADN, identifiées par la première lettre de leur nom (A, T, C, G) – déclenche la compréhension des processus génétiques par le biais de la mise en rapport analogique avec le langage. La projection métaphorique de la compréhension de l'information en tant qu'information codée dans le langage vers la compréhension – et la dénomination – du nouveau champ scientifique que représente la génétique est possible. Le modèle métaphorique du langage commence à montrer l'influence qu'il exerce sur la conception de la génétique et la dénomination des unités de compréhension. Lorsque les auteurs traitent du message que contient l'ADN, ils font appel à tout le vocabulaire existant en matière de langage et de traitement du langage. L'analogie va susciter des lexicalisations explicites.

Nous devons distinguer a) l'emprunt de termes appartenant au modèle du domaine source de l'analogie – le langage – (p. ex. *letter*) ; b) les créations de termes fondées sur le modèle du domaine source de l'analogie mais qui utilisent un mot existant dans un autre domaine (p. ex. *to sequence*) ; et c) les créations inspirées par le modèle du domaine source de l'analogie mais qui sont nommées différemment (p. ex. *codon*) (fig. 4).

| | Éléments du modèle du domaine source (le langage) utilisés dans le cadre de l'analogie | Lexicalisations relatives à l'information codée dans les gènes projetée comme un langage |
|--|--|--|
| Type a (emprunt d'un mot ou d'un terme existant) | <i>letter</i> | letter (représentation d'une base [nucléotide]) |
| Type b (mot existant recevant un sens nouveau) | Les lettres se présentent dans un ordre ou une séquence spécifique dans un texte | L'ordre des lettres peut être déterminé au moyen d'une méthode dénommée <i>sequencing</i> (séquençage) |
| Type c (création d'un nouveau terme) | Les mots représentent des unités d'information | Les codons sont des mots de trois lettres |

Figure 4. Trois types de termes de la génétique trouvant leur origine dans l'analogie avec le langage

3.2 L'ADN est un atlas

Nous présenterons en premier lieu l'analogie entre cartographie génétique et **cartographie géographique** (Jones emploie l'expression *genetic geography* (1993 : 46)) et, en second lieu, les lexicalisations que nous pouvons attribuer à ce rapport d'analogie.

L'analogie entre la représentation d'**emplacements** et de leurs distances mutuelles et les gènes a facilité la compréhension de l'importance de la position des gènes dans le génome des organismes. Chez les organismes supérieurs tels que l'homme, l'information dont on doit disposer est double : les chercheurs cherchent à savoir, d'une part, sur quel chromosome se trouve un gène donné et, d'autre part, à quel endroit de la séquence d'information génétique du brin d'ADN le gène se trouve. L'expérience du fait que l'information relative à une *terra incognita* peut être obtenue et conservée moyennant une exploration et une description cartographique minutieuse des découvertes faites a conduit les généticiens, tels des explorateurs découvrant des continents jusqu'alors inconnus, à **cartographier le génome** de différents organismes. Le génome humain est *the blueprint for the development of a single fertilized egg into a complex organism of more than 10¹³ cells. The blueprint is written in a coded message given by the sequence of nucleotide bases – the As, Cs, Gs and Ts that are strung along the DNA molecules in the*

genome (Cooper, 1994 : 71)¹. Quiconque dispose de notions élémentaires de biologie sait que l'ADN contient des gènes, que les gènes sont les messages codés utilisés pour la synthèse des protéines et que les protéines sont au cœur de toutes les fonctions d'un organisme. Il a semblé intéressant aux généticiens d'entreprendre la lecture de la séquence complète des bases d'un bout à l'autre du génome et de dresser ainsi la **carte génétique** complète de l'homme, et cela en utilisant des séquences d'ADN spécifiques, les RFLP (polymorphismes de longueur des fragments de restriction) comme **gènes marqueurs**. Cette ambition a débouché sur le projet Génome humain, un des programmes scientifiques les plus ambitieux jamais mis en œuvre. Il a pris la forme d'un effort de coopération internationale dont l'objectif était de créer une **carte détaillée de l'ADN humain**. Le projet, dont le lancement officiel remonte à octobre 1990, a vu plusieurs équipes du monde entier s'affairer à cartographier les gènes humains (dont l'estimation du nombre variait de 50 000 à 100 000 à l'époque) et à séquencer les paires de bases de ces derniers (le nombre de bases étant alors estimé à 6 milliards) (Cooper, 1994:71). L'objectif de disposer

1. NDT : [Le génome humain est] « le plan d'ensemble du développement d'un seul œuf fécondé en un organisme complexe de plus de 10¹³ cellules. Ce plan est écrit dans un message codé fourni par la séquence des bases des nucléotides – la chaîne des A, C, G et T des molécules d'ADN du génome ».

des cartes pour 2005 fut atteint en 2003. Ces cartes sont d'une valeur inestimable pour le développement de la biotechnologie, de la recherche en biologie et de la médecine clinique car elles permettent aux chercheurs de repérer l'emplacement de gènes à l'origine de maladies héréditaires sur les chromosomes.

On distingue les cartes de liaison et les cartes physiques. La différence entre les deux types de cartes peut être résumée comme suit : les cartes de liaison montrent la position d'un gène par rapport à autre tandis que les cartes physiques permettent de lire le nombre exact de paires de bases entre deux gènes. C'est la combinaison des deux types de cartes qui permet de dévoiler le génome humain.

3.2.1 La construction des cartes de liaison

Les gènes sont ordonnés de manière linéaire sur les chromosomes. Ordinairement, les gènes d'un chromosome donné sont transmis ensemble des parents aux enfants. Les gènes qui sont toujours transmis ensemble sont dits « liés ». Toutefois, un enjambement accidentel (*crossing-over*) peut se produire. Cela signifie que deux gènes qui se trouvent normalement sur le même chromosome sont transmis séparément. Les enfants qui reçoivent un chromosome modifié de la sorte d'un de leurs parents sont appelés « recombinants ». Pour établir les cartes de liaison, les chercheurs comparent le nombre de recombinants parmi les descendants de familles données et le nombre total d'enfants de ces familles. Cette comparaison fournit la fraction de recombinaison. Le principe de base qui est appliqué est que le nombre de recombinants et, partant, la fraction de recombinaison, seront d'autant plus élevés que le nombre de gènes liés est faible. En dernière analyse, les fractions de recombinaison sont converties en « **distances génétiques** » (exprimées en centimorgan), c.-à-d. en distances relatives entre gènes d'un chromosome. Ces **distances génétiques** permettent aux généticiens de déterminer l'ordre des gènes.

À l'origine, les gènes dont on testait la liaison étaient choisis de manière arbitraire (cartographie de liaison classique). Cette méthode demandant beaucoup de temps, les chercheurs ont commencé à utiliser des séquences d'ADN dites « polymorphismes de l'ADN » en tant que **marqueurs (cartographie de liaison moderne)**. Un exemple de marqueur de l'ADN est fourni par les RFLP

(polymorphismes de longueur des fragments de restriction). Il s'agit de séquences d'ADN de taille réduite provenant de la coupure de l'ADN d'un chromosome par une enzyme de restriction capable d'isoler la séquence concernée. Ces séquences, qui sont présentes chez tout être humain, peuvent être héritées sous plusieurs formes, d'où leur nom de « polymorphismes ». Une fois **localisés**, les RFLP font l'objet d'un marquage radioactif de manière à être repérables avec un microscope évolué. Enfin, on étudie la transmission conjointe des RFLP et d'autres gènes. Grâce à quoi, les chercheurs peuvent déterminer la position des autres gènes par rapport aux RFLP et entre eux.

3.2.2 La construction des cartes physiques

Les **cartes physiques** montrent les **distances physiques** (en paires de bases) séparant des **points de repère** le long d'une molécule d'ADN chromosomique. Une **carte physique** a été construite pour chacune des 23 paires de chromosomes humains. La première technique utilisée pour la construction des **cartes physiques** a été l'hybridation *in situ*. À l'instar de la technique de construction des **cartes de liaison** au moyen de RFLP, cette technique repose sur l'utilisation de sondes à ADN marquées radioactivement. En premier lieu, la double hélice de l'ADN d'un chromosome donné est séparée en deux brins simples. Ensuite, une série de séquences de paires de bases d'un des brins est colorée. Ces brins fluorescents tiendront lieu de sondes. Ils sont associés à un autre brin monocaténaire d'ADN sur lequel les généticiens vont tenter de **localiser** les séquences de bases des sondes. Les sondes s'associeront automatiquement aux séquences recherchées si celles-ci se trouvent sur le brin analysé (hybridation). Cette technique permet d'obtenir des **cartes** de basse résolution, cartes dont la caractéristique est que les marqueurs fluorescents trop proches l'un de l'autre semblent se recouvrir.

Un autre type de carte physique est celui dénommé « **carte contig** » ou simplement « contig » : *A contig map is a set of contiguous overlapping cloned fragments that have been positioned relative to one another*² (Cooper, 1994 : 113). Ces

2. NDT : Une carte contig est formée d'un ensemble de fragments clonés se chevauchant partiellement et juxtaposés les uns par rapport aux autres.

| Domaine source : carte géographique | Domaine cible : carte génétique |
|--|---|
| continents | chromosomes |
| repérer des emplacements | repérer des gènes |
| marqueurs ou points de repère | RFLP en tant que marqueurs |
| position relative d'emplacements par rapport à d'autres (en km ou <i>miles</i>) | <i>carte de liaison</i> : position de chaque gène par rapport à un autre gène (distance génétique en centimorgan) <i>carte physique</i> : nombre de paires de bases entre deux gènes (distance entre deux repères exprimée en paires de bases) |
| région | région |
| site | site |

Figure 5. Aspects de l'analogie entre cartes géographiques et cartes génétiques

cartes sont bien plus détaillées que les cartes construites au moyen de l'hybridation *in situ*. La technique employée pour les obtenir est la suivante : chaque chromosome humain est découpé en fragments (digéré) par des enzymes de restriction. Les fragments sont clonés plusieurs fois. Chacun de ces fragments est à nouveau coupé par une autre enzyme de restriction. La deuxième opération de digestion est interrompue à des moments différents pour chaque clone. Les chercheurs disposent ainsi de clones qui se chevauchent. Des sondes à ADN (obtenues par hybridation) sont utilisées pour déterminer l'ordre des séquences de bases se chevauchant dans les fragments. L'objectif final du projet Génome humain était d'obtenir des **cartes de contigs** pour la totalité du génome. Le tableau ci-dessous (fig. 5) montre l'analogie entre le domaine source (carte géographique) et le domaine cible (carte génétique) du modèle du sous-domaine métaphorique.

3.3 L'ADN est un programme informatique

Dans le cas du troisième sous-modèle de l'information, le matériel génétique (ADN) peut être compris comme un programme d'un système informatique (la cellule). Cela ressort de manière explicite dans les extraits suivants : *There's another appropriate analogy here, one that emphasizes why the unity of life has been so important to molecular biology. You can think of DNA's coded instructions as molecular "software" that*

runs the "hardware" of life. Just as a word-processing program tells computer hardware what to do, DNA's instructions control life's machinery. Why is that comparison useful? Because if you work with computers the way most of us do, you know enough about your favorite word processor or spreadsheet to use it, but you could hardly write the program yourself. In much the same manner, molecular biologists know enough about certain DNA-based "programs" to use them without fully understanding how they work.

So it's handy that much of the life's software – regardless of the organism it comes from – will run on the hardware of nearly any other living cell. That's why, for example, researchers who discover and learn to control the molecular word processor used by one organism can harness that tool to manipulate genetic text in different organisms without having to learn precisely why or how that particular molecular program works as it does (Levine et Suzuki, 1993 : 23).

As pointed out above, the subunits of the DNA strands, the nucleotides, are the chemical basis for storage of information in DNA (Drlica, 1992 : 33).

Les « instructions codées de l'ADN » (*DNA's coded instructions*) sont conçues comme le « logiciel moléculaire exécuté par le matériel de la vie » (*molecular software that runs the hardware of life*). La comparaison est opératoire parce que, de la même manière que les utilisateurs d'un ordinateur sont capables d'utiliser un programme sans être des programmeurs,

3. NDA : nous surlignons en caractères gras.

les biologistes moléculaires disposent de connaissances suffisantes sur l'ADN pour pouvoir l'utiliser sans devoir en comprendre complètement le fonctionnement. Le « logiciel » (ADN) d'un organisme spécifique peut être exécuté sur le matériel (les cellules) de quasiment tout autre organisme.

Notre corpus fait bien mention de l'analogie entre un logiciel informatique et l'ADN mais nous n'y avons trouvé aucun exemple qui illustrerait la manière dont cette analogie a inspiré la formation de termes.

3.4 L'ADN est un film

Le modèle de l'information comporte, en puissance, d'autres sous-modèles que les trois sous-modèles mentionnés jusqu'à présent (langage, géographie et logiciel informatique). Dans Drlica (1984 et 1992), nous constatons la mise en avant d'une quatrième analogie (fig. 6) pour expliquer la structure de l'information génétique.

In some ways DNA is similar to motion picture film. Like film, DNA is subdivided into "frames" that make sense when seen in the correct order. In DNA the "frames" correspond to the letters in the genetic code, { ... }. When a number of frames or genetic letters are organized into a specific combination, they create a scene in the case of film and a gene in the case of DNA (Drlica, 1992 : 4).

As pointed out above, the subunits of the DNA strands, the nucleotides, are the chemical basis for storage of information in DNA. Returning to the film analogy { ... }, the units we have now defined as nucleotide pairs, or base pairs, correspond to a scene in the motion picture film (Drlica, 1992 : 34).

| | |
|------------------------|---------------------------|
| <i>Film (cinéma)</i> | ADN |
| <i>frames (plans)</i> | lettres du code génétique |
| <i>scenes (scènes)</i> | gènes |

Figure 6. L'analogie entre un film et l'ADN

Les lexicalisations *frames* et *scenes* sont idiolectales et employées dans le cas de Drlica à des fins didactiques. Elles sont temporaires. Les lexèmes n'acquièrent pas le statut de termes de la langue des sciences biologiques. Dans l'ouvrage de Drlica, cette analogie vient renforcer la compréhension du concept de *gene splicing* (épissage des gènes). Les gènes peuvent être épissés en laboratoire à la manière dont un film ou une bande de pellicule peut être édité dans un

studio de cinéma (le terme *splicing* est abordé en détail dans Temmerman (1998), chapitre 6 et Temmerman (à paraître)).

4 Métaphore créative et métaphore didactique

Les citations présentées dans la deuxième partie de notre article, extraites de textes rédigés à destination des chercheurs ou à des fins de vulgarisation, nous ont permis de démontrer qu'un modèle métaphorique a une structure interne (les sous-domaines). Certains sous-domaines sont exploités pour faciliter la compréhension (tel le sous-domaine « film »), d'autres ont pour finalité d'« inventer » une nouvelle compréhension scientifique. Il convient de faire la différence entre les articles des chercheurs dans lesquels l'analogie et les néologismes lexicalisés sont le fruit d'une conceptualisation analogique consciente ou subconsciente, les manuels spécialisés pour un lectorat formé aux sciences, dans lesquels l'analogie se retrouve, et les textes de vulgarisation rédigés par des spécialistes pour un vaste lectorat, dans lesquels l'analogie est rendue plus explicite à des fins didactiques (fig. 7).

| | |
|--|-----------------------------|
| Type de texte 1 : articles scientifiques | introduction de l'analogie |
| Type de texte 2 : manuels spécialisés | présence de l'analogie |
| Type de texte 3 : textes de vulgarisation | explicitation de l'analogie |

Figure 7. Déploiement de l'analogie dans les différents types de textes

5 Traduire la terminologie métaphorique des sciences de la vie

Dans les parties précédentes, nous avons montré que la pensée analogique et les néolexicalisations métaphoriques jouent un rôle important dans la langue des sciences de la vie. Jusqu'à présent, nous n'avons pris en considération que

la langue anglaise. Deux questions pertinentes se posent pour les traducteurs de la littérature relative aux sciences de la vie : ces analogies et les lexicalisations qui s'ensuivent sont-elles universelles ? et comment traduire la terminologie métaphorique de l'anglais dans d'autres langues ?

En ce qui concerne les langues sur lesquelles nous estimons pouvoir nous prononcer – à savoir les langues romanes et germaniques du fait que nous en possédons une maîtrise suffisante pour pouvoir juger des textes traitant des sciences de la vie et les langues slaves du fait de nos contacts avec des personnes qui en ont une bonne maîtrise –, nous pouvons affirmer sans risque

que ces langues et la culture qui les sous-tend ont accès au domaine source pour les quatre sous-modèles de l'information distingués en deuxième partie de notre article. Il serait intéressant d'investiguer de quelle manière des langues sans système d'écriture de type phonologique (tel le chinois) gèrent le rapport d'analogie entre les quatre lettres et les quatre bases.

La figure 8 montre des exemples d'équivalents en français et en néerlandais de la terminologie métaphorique anglaise. Nous pouvons observer l'équivalence intégrale des traductions littérales, des calques et des emprunts.

| Anglais | Français | Néerlandais |
|---------------|-----------------|-------------------------|
| genetic code | code génétique | genetische code |
| transcription | transcription | transcriptie |
| translation | traduction | translatie/vertaling |
| genetic map | carte génétique | genetische kaart |
| to map | cartographier | mappen/in kaart brengen |
| sequencing | séquençage | sequencen/sequeneren |

Figure 8. Quelques néologismes métaphoriques en anglais et leurs équivalents en français et en néerlandais

Les futurs traducteurs et terminologues devraient être initiés au phénomène de la modélisation analogique de la pensée et à son influence sur les processus de lexicalisation à l'œuvre dans le cadre du développement de nouvelles orientations des disciplines scientifiques comme les sciences de la vie. Dans Temmerman (2000[b]) et Temmerman (2002), nous montrons comment, dans le cadre de la formation à la terminologie sociocognitive, de futurs terminographes sont initiés à plusieurs méthodes d'analyse textuelle. Une de ces méthodes met l'accent sur la détection de modèles de pensée analogique sous-jacents qui laissent leur trace dans la terminologie métaphorique. Les exemples commentés dans cet article devraient démontrer que la qualité des traductions faites par des traducteurs qui maîtrisent le modèle métaphorique sous-jacent aux sciences de la vie sera probablement meilleure.

6 L'intégration sociale et culturelle de la cognition

Sharifian (2003) souligne que les membres d'un groupe culturel négocient et renégocient leur cognition culturelle dans le temps et l'espace. Cette vision de la cognition fournit une base pour la compréhension de la conceptualisation expérientielle au niveau des groupes. Les conceptualisations culturelles sont émergentes dans le sens où elles émergent entre les membres d'un groupe culturel et elles sont réparties dans la mesure où elles ne sont « pas imprimées de manière égale dans l'esprit de chaque membre » (Sharifian, 2003 : 196). Dans cette perspective, le langage est considéré à la fois comme « un système de cognition répartie » (Sharifian, 2003 : 196) et comme un système « qui incarne dans une vaste mesure les conceptualisations culturelles de l'expérience de ceux qui les énoncent » (Sharifian,

2003 : 200). Sharifian préfère le terme de *conceptualisation* à celui de *concept* car il reflète la nature dynamique de la cognition. Il distingue différents types de conceptualisations telles la schématisation, la catégorisation, la métaphore et les « intégrations conceptuelles » (*conceptual blends*). Le langage est conçu autant comme un système réparti que comme un gisement de la conceptualisation culturelle.

Les traducteurs devraient être formés à la compréhension de la cognition saillante dans les termes et les descriptions des publications scientifiques (qu'il s'agisse d'articles scientifiques en tant que tels ou de littérature de vulgarisation de ces derniers). Dans Temmerman (2000[a], chap. 5) et Temmerman (à paraître), nous avons montré que la sensibilité à la pensée analogique et à la métaphore a un rôle important à jouer dans le développement des terminologies nouvelles des sciences de la vie. Nous y relevons aussi comment les changements d'ordre social, cognitif et technologique s'accompagnent d'évolutions de la langue et, pour les sciences de la vie, comment l'étude d'archives textuelles de ces sciences permet une compréhension plus profonde de ces quatre processus parallèles et de leur influence réciproque.

Dans Temmerman (à paraître), nous décrivons la nature complexe de la catégorie du terme *splicing*. Nous présentons la structure prototypique de *splicing* en analysant de quelle manière la flexibilité de cet élément lexical à la structure prototypique résulte dans une large mesure d'un mécanisme de polysémisation ainsi qu'en étudiant le rôle que les modèles métaphoriques ont joué dans le développement de la catégorie.

La prise en considération du système qui sous-tend le tout n'est pas sans conséquence pour les descriptions terminologiques (plurilingues) et la représentation des connaissances. Reconnaître la force créatrice du système de la langue compte tenu du caractère socialement et culturellement situé de cette dernière peut avoir un impact sur la gestion de ressources terminologiques dynamiques capables de rendre compte de l'évolution du sens (Temmerman, 2002 ; Temmerman et Kerremans, 2003).

Rita Temmerman,
Erasmushogeschool Brussels,
Centrum voor Vaktaal en
Communicatie,
Department of Applied
Linguistics,
Bruxelles,
rita.temmerman@ebb.be.

Traduit de l'anglais par France
et Paul Muraille-Dombret.

Bibliographie

- Alberts (B.), Bray (D.), Lewis (J.), Raff (M.), Roberts (K.) et Watson (J.), 1994 : *Molecular biology of the cell*, New York et Londres : Garland Publishing.
- Berg (P.) et Singer (M.), 1992 : *The Language of Heredity*, California : Blackwell Scientific Publications.
- Cooper (N.), 1994 : *The Human Genome Project*, California : University Science Books.
- Drlica (K.), 1984 : *Understanding DNA and Gene Cloning*, New York : John Wiley and Sons.
- Drlica (K.), 1992 : *Understanding DNA and Gene Cloning*, 2^e éd., New York : John Wiley and Sons.
- Hayes (B.), 1998 : “The Invention of the Genetic Code” dans *American Scientist*, n° 1.
- Hodson (A.), 1992 : *Essential Genetics*, London : Bloomsbury.
- Lakoff (G.) et Johnson (M.), 1980 : *Metaphors we Live By*, Chicago : University of Chicago Press.
- Lakoff (G.), 1987 : *Women, Fire and Dangerous Things*, Chicago : University of Chicago Press.
- Levine (J.) et Suzuki (D.), 1993 : *The Secret of Life*, Boston : WGBH Educational Foundation.
- Nicholl (D.), 1994 : *An Introduction to Genetic Engineering*, Cambridge : Cambridge University Press.
- Shapiro (R.), 1991 : *The Human Blueprint*, New York : St. Martin's Press.
- Sharifian (F.), 2007 : “On cultural conceptualisations” dans *Journal of Cognition and Culture*, n° 3 (3), p. 187–207.
- Temmerman (R.), 1998 : *Terminology Beyond Standardisation. Language and Categorisation in the Life Sciences*, Leuven Catholic University (thèse de doctorat sous la direction du prof. Dirk Geeraerts).
- Temmerman (R.), 2000[a] : *Towards New Ways of Terminology Description: the Sociocognitive Approach*, Amsterdam/Philadelphia : John Benjamins.
- Temmerman (R.), 2000[b] : “Training Terminographers: The Sociocognitive Approach” dans *Proceedings Euralex Stuttgart*.
- Temmerman (R.), 2000[c] : « Une théorie réaliste de la terminologie : le sociocognitivism » dans *Terminologies Nouvelles*, n° 21, p. 58-64.
- Temmerman (R.), 2002 : “Metaphorical models and the translation of scientific texts” dans *Linguistica Antverpiensia*, n° 1, p. 211–226.
- Temmerman (R.) et Kerremans, (K.), 2003 : “Termonography: Ontology building and the sociocognitive approach to terminology description” dans *Proceedings of the XVII International Congress of Linguistics (24–29 July 2003)*, sur cédérom, Prague : Universita Karlova.
- Temmerman (R.), à paraître : “Sociocultural situatedness of terminology in the life sciences: The history of splicing.” dans Dirven (Fr., Ro. et Re.), Zlatev (J.) et Ziemke (T.), *Body, Language and Mind. Vol II. Interrelations between Biology, Linguistics and Culture*, Berlin/ New York : Mouton de Gruyter.
- Site de l'Indiana State University, http://web.indstate.edu/thcme/mwking/protein-synthesis.html#genetic_code