

---

## Notes de lecture

Rubrique préparée par Denis Maurel

*Université de Tours, LIFAT (Laboratoire d'informatique fondamentale et appliquée)*

---

**Elisabetta JEŽEK. The Lexicon: An Introduction (Oxford Textbooks in Linguistics). Oxford University Press. 2016. 234 pages. ISBN 978-0-19960-153-0.**

Lu par **Kevin BRETONNEL COHEN**

*University of Colorado School of Medicine, Computational Bioscience Program, Biomedical Text Mining Group – Université Paris-Saclay*

---

*Nous nous trouvons en pleine recrudescence d'une implication de la sémantique dans le domaine du TAL. C'est un essor inouï depuis l'ère des heurts chomskiens-schankiens. Dans un certain sens, cette résurgence de la sémantique dans un contexte de dominance de l'apprentissage machine s'est amorcée avec la naissance de l'étiquetage de rôles sémantiques chez Gildea et Jurafsky, et bat son plein depuis l'avènement du trio PropBank, VerbNet et SemLink. Cet essor s'accélère encore plus vivement avec la renaissance des réseaux neuronaux, l'apprentissage profond et les plongements lexicaux.*

La sémantique, de nos jours, se fonde sur une démarche fort différente des précédentes. De la sémantique dite « non profonde » de PropBank à la sémantique proprement dite « profonde » de VerbNet et SemLink, ces prédécesseurs dans notre champ d'enquête se fondent sur des suppositions rationnelles, à savoir sur les socles de la logique, de l'implication et des rôles thématiques. C'est une approche de la représentation du sens qui a comme source un patrimoine millénaire, qui a connu un succès d'une grande portée autant dans les disciplines langagières expérimentales (apprentissage des langues chez les enfants, psycholinguistique...) que dans les disciplines formelles.

Au lieu de la démarche des représentations formelles, les approches fondées sur l'apprentissage profond et les plongements lexicaux sont fortement reliées à une version de la théorie distributionnelle de Firth, ainsi que de celle de Harris. Or, on pourrait dire que la sémantique des plongements lexicaux s'appuie sur une version de la sémantique distributionnelle réduite, voire affaiblie. Tout d'abord, les sémantiques distributionnelles, à l'opposé des plongements lexicaux, ne se restreignent pas aux lexèmes. Loin de cela, la théorie des répartitions lexicales, selon Harris, est explicitement et inéluctablement en lien avec les classes sémantiques, celles-ci étant explicitement définies en tant que telles. La sémantique distributionnelle de Firth s'étend bien au-delà de ces classes sémantiques, allant jusqu'aux contextes sociaux, sémiotiques, etc. On pourrait dire que la sémantique dite « distributionnelle » des plongements lexicaux oublie à la fois ses aïeux, les distributionnalistes des années

1950-1970, et ses ancêtres plus lointains, les formalistes. Un rapprochement entre les deux approches est autant possible que désirable. Ce qui nous manque, c'est une ressource pour approfondir la connaissance de la sémantique des composants des plongements lexicaux et autres représentations distributionnelles.

Elisabetta Ježek comble à merveille ce fossé entre la sémantique distributionnelle de nos jours, la sémantique formelle de PropBank, VerbNet et SemLink, la sémantique distributionnelle de Firth et Harris, et les millénaires de théorisation antécédents. Le livre débute avec les composants du sens, soit la problématique de la définition du « mot », l'information lexicale, et le sens lexical. Puis on se tourne vers la structure du lexique, à savoir les structures paradigmatiques et les structures syntagmatiques. Ces dernières créent un pont entre les représentations formelles et les modèles statistiques tels que les collocations. En bref, cet ouvrage traite du lexique dans son sens large, à savoir de sa structure globale, de sa structure paradigmatique, de son lien avec les rapports syntagmatiques et des problématiques théoriques sous-jacentes desdits sujets, ainsi que des approches concrètes vers celles-ci. Dans son contexte plus étendu relatif à d'autres ouvrages dans le domaine de la sémantique lexicale en traitement automatique de langage naturel, ce livre fournit un pivot précieux entre les approches distributionnelles et les approches s'appuyant sur la structure prédicat-argument, ainsi que sur la psycholinguistique, et les ouvrages de sémantique au sens large (à savoir, la sémantique de la phrase, le sens en pragmatique, etc.). Le contenu est présenté très clairement et de façon à ce qu'il soit accessible à tous.

Le premier chapitre introduit les idées de base du lexique, telles que le contraste entre lexique, dictionnaire, et vocabulaire, le lien entre lexique et sémantique, le contraste entre sens lexical et sens grammatical, et surtout la notion du « mot ». Tandis qu'en traitement automatique de langage, on fait de plus en plus la part belle au mot, surtout depuis l'avènement des plongements lexicaux, cette notion du mot, y compris sa circonscription voire son statut ontologique, reste encore contestée, bien qu'elle ait toujours été (et reste encore) sujette à polémique chez les linguistes. On peut dire que ce problème devient de plus en plus important avec l'intérêt accru pour la traduction automatique. En lien avec cette problématique de la circonscription et de la définition du « mot », ce chapitre propose les pistes les plus précieuses et pratiques, tels l'ordre immuable et la cohésion.

Le deuxième chapitre aborde le contenu du lexique. Les données le constituant sont divisées en informations lexicales au sens propre, et informations dites encyclopédiques ou connaissances du monde. La première catégorie exige une distinction entre différents types de sens, notamment la dénotation d'un mot, sa capacité à circonscrire une certaine classe, et ses connotations ou ses associations affectives et de registre. On ajoute à ces deux genres de sens l'acception collocationnelle, le cas échéant. Au-delà de ces genres de signification, l'information lexicale inclut des renseignements sur la forme, que ce soit phonologique ou typographique, et sur la morphologie, cette dernière se trouvant à cheval entre le seuil morphologie et syntaxe. On y rencontre ainsi les parties du discours, la morphologie flexionnelle et dérivationnelle, et concernant les prédicats, le schéma actanciel et, cas exceptionnel, l'Aktionsart.

Ces genres d'informations lexicales ayant été élaborées, ce chapitre s'oriente vers la distinction entre l'information lexicale et la connaissance encyclopédique. Sans prise de position sur cette problématique polémique, ce chapitre expose une gamme de positions : minimaliste, intermédiaire, maximaliste, et inexistante (c'est-à-dire, la position selon laquelle faire une distinction lexicale et encyclopédique n'est pas cohérent).

Le troisième chapitre est consacré au sens. Tandis que les chapitres suivants sont dédiés aux aspects pertinents du lexique. Pour ceux qui s'occupent de la conception de ressources lexicales, ce chapitre est une réflexion bien approfondie sur la définition des acceptions situées au noyau des dites ressources et sur les implications des théories du sens sous-jacentes. De surcroît, c'est dans ce chapitre que l'on apprécie les liens entre le lexique et plusieurs autres sujets de recherche.

D'abord, il traite des caractères des mots qui se révèlent presque toujours les plus problématiques dans le traitement automatique du langage, à savoir, l'ambiguïté et la polysémie. Les cas d'ambiguïtés les plus extrêmes, pourtant les plus faciles à désambigüiser, comme en anglais *pole* (endroit, *magnetic pole*) et *pole* (perche, *fiberglass pole*), ne sont guère traités. En revanche, les cas d'ambiguïtés les plus répandus et (plus) difficiles, à savoir la polysémie des mots possédant plusieurs acceptions étroitement liées, sont bien élaborés.

À partir de là, on se dirige vers plusieurs théories du sens. On peut classer ces théories comme « internes » et « relationnelles ».

Parmi les premières, on trouve la théorie appelée référentielle. Selon cette théorie les mots sont utilisés en tant que choses renvoyant à des objets et à des événements qui existent ou qui se passent dans le monde. Cette thèse fait la part belle au monde dans lequel le langage se trouve. Parmi les avantages de cette thèse, on doit compter le fait qu'elle fait partie intégrale de la sémantique dite formelle. Quant aux inconvénients, on constate qu'ils sont difficiles à incorporer dans une approche sur le traitement automatique des langues fréquentiste à cause de son déterminisme.

À cette base référentialiste, l'hypothèse dite « mentaliste » ajoute la notion du « concept », répondant ainsi à une faiblesse de la théorie référentielle, le fait que l'on peut renvoyer à des choses sans existence, par exemple des abstractions non matérielles (*happiness, beauty*) ou non existantes (*unicorns*). À l'opposé du monde extérieur au sein de la référentialité au sens pur, le mentalisme fait la part belle à l'individu. Cette approche vis-à-vis du sens bénéficie d'un lien avec les ontologies qui font partie intégrale des approches computationnelles par rapport à la résolution de la coréférence et à l'informatique médicale fournissant couramment beaucoup de cas d'utilisation du traitement automatique des langues. En revanche, en pratique, elle est limitée par les mêmes problèmes de couverture qui bloquent souvent l'utilisation des ontologies dans des applications du traitement automatique des langues.

Les deux théories précédentes, référentielle et mentaliste, ont en commun des catégories qui sont forcément conscrites. Ceci est un problème pour une conception de la signification, car, de toute évidence, les catégories des concepts n'ont pas de fron-

tières bien définies, mais plutôt des limites assez floues. C'est le cas soit chez les locuteurs d'une langue, soit au sein du lexique, où l'on croise beaucoup d'acceptions connexes subtilement différentes ou souvent difficiles à différencier, même en contexte (*university* en tant qu'organisme, en tant que ses salariés et ses élèves, et en tant que lieu).

La théorie dite *du prototype* émet une conception très différente du concept structuré, précisément une structure avec un noyau – le *prototype* et des composants plus ou moins lointains de ce prototype. Par exemple, le prototype du *concept bird* serait (pour un anglophone) *robin* (rouge-gorge), avec un composant éloigné du noyau *vulture* (vautour) et des composants encore plus lointains *ostrich* (autruche) et *penguin* (manchot). Cette conception plus relationnelle qu'interne de la signification a plusieurs avantages, notamment l'introduction de la similarité à la variabilité des acceptions des mots. Elle apporte aussi l'idée de similarité à la signification, ce qui se révèle primordial pour les approches distributionnelles si courantes de nos jours. On peut donc suivre les traces de la signification depuis les théories relationnelles jusqu'à Aristote, bien qu'il reste à savoir comment concevoir des mises en œuvre informatiques concrètes qui peuvent se servir des mérites des approches aux deux extrémités de cette gamme de possibilités.

Le quatrième chapitre traite du lexique en tant que « structure », à savoir un ensemble organisé de mots connexes. On parle de deux types d'organisation au sein du lexique : le regroupement en classes et les relations du sens. Par « classes » on entend des ensembles de mots regroupés en fonction de comportements morphologiques ou syntaxiques. Un des mérites de ce chapitre est de prendre en considération des idées que l'on rejette parfois facilement, et peut-être trop facilement, en particulier une réflexion approfondie sur les liens entre les parties du discours (*word classes*) et le sens. Un long discours sur la notion de la « catégorie ontologique » de Lyons et sur la « stabilité temporelle » de Givón relie les perspectives linguistiques et morphosyntaxiques aux perspectives philosophiques courantes par le biais des concepts d'endurants, de perdurants, de continuants et d'occurants. Ce discours prolongé se révélera très utile pour tous ceux qui côtoient des ontologues. Ce chapitre prépare la voie qui mène éventuellement au lien avec les représentations distributionnelles du sens, non pas par le biais de la représentation des liens, mais plutôt par celui de la perspective des exigences des mots (telles la structure actancielle (argument structure) et la structure d'événements).

Le cinquième chapitre aborde les structures paradigmatiques au sein du lexique. Cette notion de paradigme comprend les mots qui « rivalisent » entre eux, par exemple synonymes et antonymes, bref, des mots qui peuvent occuper la même position dans une phrase. On croise alors des relations autorisant des implications, telles que l'hyponymie et l'hyponymie (*vehicle* et *car*, *move* et *walk*) et la méronymie et l'holonymie (*finger* et *hand*, *athlete* et *team*), ainsi que les complexités de la synonymie et les relations d'opposition. À partir de cela, on se tourne vers les relations de causalité, de finalité, et d'implication (sens *entailment*). Pour les linguistes, ce sont des relations très intéressantes à cause de leur utilité dans l'enquête des sèmes premiers (sémantiques primitives).

Le sixième chapitre est consacré aux relations dites « syntagmatiques ». On aborde ce sujet dans un livre traitant du lexique, d'un ensemble de mots en tant qu'individus, puisque certains mots montrent une répartition limitée par rapport à d'autres mots. Cette répartition étant un caractère d'un mot spécifique, c'est dans le lexique d'une langue que cela est spécifié, bien qu'il s'agisse d'un aspect combinatoire du mot.

Globalement on peut taxonomiser les mots en ce qui concerne leurs contraintes combinatoires en suivant plusieurs axes. Il y a une dimension concernant ce que l'on pourrait qualifier de source des contraintes : des limites probabilistes d'origine conventionnelle, et des limites proprement dites sémantiques ou ontologiques. De surcroît, il y a la présence ou l'absence d'interprétation compositionnelle d'une combinaison de mots. En outre, on peut classer une contrainte par la nature des résultats de sa violation : un énoncé agrammatical ou un énoncé sémantiquement anormal. On peut aussi différencier les combinaisons par la mobilité ou la raideur des relations syntaxiques entre les mots liés. Enfin, on est amené à la conclusion que, contre les affirmations de certains, on ne peut pas considérer les combinaisons de mots (et leurs contraintes) comme un continuum des combinaisons libres aux idiotismes.

Somme toute, on peut conseiller ce livre, agrémenté d'exemples aussi bien nombreux que clairs, à tous ceux qui s'occupent de la conception et de la construction des ressources lexicales et/ou s'intéressent à la sémantique en traitement automatique des langues naturelles ou à la sémantique lexicale, voire générale.

---

**Frédéric LANDRAGIN. Comment parler à un alien? *Le Béliat éditions*. 2018. 263 pages. ISBN 978-2-84344-943-7.**

Lu par **Yannis HARALAMBOUS**

*IMT Atlantique et UMR CNRS 6285 Lab-STICC*

---

*Frédéric Landragin, directeur de recherche au CNRS, membre du laboratoire LATTICE et spécialiste, entre autres, du dialogue homme-machine multimodal, nous offre ici une introduction moderne, claire et pédagogique à la linguistique, enrichie par sa grande culture et par sa passion pour la science-fiction. Cet ouvrage se caractérise par la générosité de son contenu dans les deux domaines (linguistique et science-fiction), où rien (ou presque) n'a été laissé de côté.*

Une question que l'on peut se poser est celle du public ciblé. Je peux dire avec certitude qu'il y aura au moins deux types de lecteurs : (a) les amateurs de science-fiction éclairés (c'est-à-dire dont la vision de ce genre littéraire ou cinématographique dépasse les sabres laser de Luke Skywalker et le bikini doré de la princesse Léia) qui y trouveront une véritable introduction à la linguistique, introduction qui ne cessera pas de les motiver en leur faisant (re)découvrir des textes de science-fiction emblématiques, et (b) les linguistes en quête d'une introduction à leur discipline accessible au grand public, et plus moderne qu'*Alice au pays du langage* de Marina Yaguello (pour ma part, j'ai fait acheter plusieurs exemplaires de

l'ouvrage de Landragin à la bibliothèque de mon institution puisque je vais m'en servir pour mon cours de TAL). Pourrait-on imaginer un lectorat qui ne s'y connaisse ni en linguistique ni en science-fiction et qui découvrirait ainsi les deux domaines ? Le temps le montrera, et j'imagine que cela fait partie du pari de l'auteur. En tout cas, le choix de la maison d'édition Le Bélial est judicieux : il s'agit d'un « petit » éditeur de science-fiction (« petit » comparé aux éditions du Seuil dont la collection « Folio SF » domine le marché), éditeur qui anime un forum de discussion francophone très actif (870 membres, plus de 60 000 messages) et publie une riche revue trimestrielle (*Bifrost*) depuis plus de vingt ans.

L'ouvrage comporte une introduction et cinq chapitres assez variés. Si le premier chapitre satisfera aussi bien les linguistes que les amis de la science-fiction, le deuxième et le troisième chapitre portent plus sur la linguistique, le quatrième est un excellent mélange de science-fiction et de linguistique, et le cinquième porte plus sur la communication hypothétique avec les aliens, en accord avec le titre de l'ouvrage.

L'introduction, véritable chapitre en soi, fournit les notions fondamentales : on y trouve la différence entre *langue* et *langage* (tout en notant au passage que le titre du roman bien connu *Les langages de Pao* est erroné), les définitions de *signifiant* et *signifié* selon Saussure, la *double articulation*, les *fonctions du langage* selon Jakobson, la *théorie des actes du langage*, ce qu'est et ce que n'est pas la linguistique, etc. Entre deux notions linguistiques, l'auteur réussit à placer des références à des ouvrages ou films de science-fiction : cette alternance thématique permet de mettre en perspective les notions afin de mieux « faire avaler la pilule » du caractère théorique des notions.

Le premier chapitre se focalise sur le sous-genre de la science-fiction qui fait intervenir la linguistique. Au départ, tout en mentionnant un bon nombre d'œuvres connues et moins connues, l'auteur se concentre surtout sur le roman *L'Enchâssement* de Ian Watson, ce qui lui permet d'introduire le côté génératif de la théorie syntaxique de Chomsky. Il enchaîne avec l'hypothèse de Sapir-Whorf, et l'antagonisme entre celle-ci et l'approche chomskyenne, en incluant des exemples d'œuvres de science-fiction dans un camp ou dans l'autre. Il conclut le chapitre avec un bref aperçu de quelques grands classiques de la science-fiction (*Nous autres* de Zamiatine, *1984* d'Orwell, *Babel 17* de Delany, *Légationville* de Miéville) à la lumière des notions étudiées.

Le deuxième chapitre traite de la vision diachronique des langues : leur naissance et leur évolution, leur diversification, leurs familles. On y apprend que le sens des mots évolue, que dans chaque communauté linguistique il y a toujours ceux qui veulent figer une langue et ceux qui la font vivre par un perpétuel changement. Landragin mentionne le mythe raciste du XIX<sup>e</sup> siècle qui voulait les langues flexionnelles (comme la nôtre, pardi !) plus évoluées que les agglutinantes et les isolantes, mythe utile à rappeler non pas pour son intérêt historique, mais parce qu'un certain nombre d'œuvres de science-fiction s'en servent. Puis, il s'intéresse au futur, avec une section sur l'anticipation linguistique qui permet de distinguer les œuvres sérieuses et réfléchies de science-fiction des produits commerciaux du type

*La planète des singes* (je parle du film et non pas du livre de notre compatriote Pierre Boulle !) où, deux mille ans après notre ère, une espèce simienne a pris le contrôle de la planète et, comme par hasard, parle... un anglais parfait. Qui dit anticipation linguistique dit anticipation historique et cela permet à Landragin de clore le chapitre par quelques éléments de sociolinguistique, toujours illustrés par des exemples d'œuvres de science-fiction paradigmatiques.

Avant de passer au troisième chapitre, je ne peux m'empêcher de mentionner ce qui constitue, à mon humble avis, le seul « couac » de cet excellent ouvrage : son attachement à la conception phonocentriste. En effet, Landragin affirme que « *la langue est orale et l'écriture n'en est qu'une représentation graphique* ». Or, depuis les années 40 des auteurs comme Vachek, Hjelmslev, Uldall, Anis, Catach, Sproat, Coulmas, Dürscheid, Neef et Sébastianoff ont redonné à la modalité écrite de la langue ses lettres de noblesse, et selon eux, celle-ci possède, bel et bien, une double articulation. Il existe une discipline dédiée qui est le pendant de la phonologie : la « graphématique ». Et en Allemagne on parle même de « grapholinguistique » (*Schriftlinguistik*) puisque la modalité écrite touche tous les niveaux d'étude de la langue. Je trouve cela bien réducteur de revenir aux dogmes saussuriens, mais cela bien sûr n'engage que moi.

Dans la lancée du chapitre sur la vie des langues naturelles, le troisième chapitre traite de langues artificielles (thème que l'on retrouve dans l'excellent *Les langues imaginaires* de Marina Yaguello, ainsi que dans *La recherche de la langue parfaite* d'Umberto Eco). Après le passage obligé sur les langues auxiliaires internationales (espéranto et volapük) et une brève section sur les « langues fondées sur la logique » (illustrées par *The Troika Incident* de Brown), l'auteur enchaîne avec les langues fictionnelles, en fournissant un grand nombre d'exemples allant du XIV<sup>e</sup> siècle jusqu'à la langue dothraki du bien actuel *Game of Thrones*. Mais il ne s'arrête pas là : en effet, il fait la (très subtile) distinction entre langues fictionnelles représentant des langues naturelles et langues fictionnelles représentant des langues artificielles, et il enchaîne avec une autre série d'exemples, tout aussi intéressants, du deuxième cas. Une brève mention de la *glossolalie* (c'est-à-dire l'utilisation d'une soi-disant langue inconnue) permet d'introduire le critère des *hapax* de Marina Yaguello : la distribution des langues naturelles est telle qu'on y trouve entre 46 % et 48 % d'*hapax*, le cas de glossolalie le plus connu (Helene Smith, 1861-1929) ne comportait que 32 % d'*hapax*, donc il s'agissait très probablement d'une escroquerie.

Suit une section dédiée à la morphologie dérivationnelle, aux néologismes et, plus généralement, aux différents procédés de variation ou de création lexicale adoptés dans la science-fiction. L'auteur donne des conseils aux futurs auteurs de science-fiction sur les pièges à éviter en jouant sur le lexique et la morphologie. La dernière section de ce chapitre est très intéressante puisque l'on entre dans le domaine de l'analyse du discours de la science-fiction, notamment en ce qui concerne la « distanciation cognitive » de Darko Suvin et la « xéno-encyclopédie » d'Irène Langlet, qui caractérisent le processus de lecture de la science-fiction moderne. Cette section donne vraiment envie de se plonger dans les travaux d'analyse des procédés spécifiques à la science-fiction.

Je note au passage que si j'étais l'auteur de ce livre, j'aurais peut-être insisté un peu plus sur la notion de langage contrôlé, sur le fait qu'il existe un véritable continuum entre les langages formels et les langages naturels – et même un indicateur à cinq dimensions pour positionner un langage dans cet espace, donné par Tobias Kuhn –, et sur les tentatives de création de langages de programmation « proches du langage naturel » comme *Inform 7*. D'ailleurs l'auteur ne mentionne qu'une seule fois les langages de programmation dans un paragraphe de dix-sept lignes citant un texte vieux d'une cinquantaine d'années...

Partie centrale de l'ouvrage, le chapitre 4 est particulier et passionnant à lire : l'auteur prend une à une toutes les couches d'étude de la langue (à l'exception, bien sûr, de la graphématique !) et les décrit en s'inspirant d'œuvres de science-fiction. Il commence par le niveau lexical avec quelques notions de terminologie et quelques exemples de structuration de champs lexicaux différant selon les langues (illustrations : les *Borogoves* de Kuttner et Moore, traduit par Boris Vian !, et *LAMA* d'Egan). Puis il passe au niveau phonétique, où il parle évidemment de phonèmes et de l'alphabet phonétique international, mais aussi de langues sifflées. La phonétique est illustrée par l'angoissant *Épépé* de Karinthy où un linguiste se trouve par erreur dans un pays dont il ne comprendra, jusqu'à la dernière page du livre, pas un seul mot de la langue. Une section est dédiée à la prosodie, illustrée par le célèbre « *Je suis désolé, Dave* » énoncé par l'ordinateur HAL (acronyme décalé d'une lettre d'« IBM ») de *2001, Odyssée de l'espace* de Clarke, ordinateur qui fait toujours froid dans le dos aujourd'hui malgré ses cinquante ans bien sonnés.

Suit le niveau morphologique qui, contre toute attente, est fascinant puisqu'on y découvre le *passé antérieur surcomposé de subjonctif futur semi-conditionnel plagal 2<sup>e</sup> forme* de Douglas Adams, les *préfixes et suffixes* de 1984, les déboires d'un prêtre qui s'est fait mutiler à cause de la mauvaise interprétation d'une dérivation morphologique, les *morphèmes évidentiels* des langues amazoniennes, et le fait que la novlangue de ce même 1984 permet, malgré sa rigoureuse planification, de former des phrases volontairement ambiguës, comme la bien connue « la jeune porte le voile ».

L'exemple qui illustre le niveau syntaxique, tiré de la *Guerre des étoiles*, « dans le bon ordre les mots placer tu dois » est immédiatement reconnaissable comme pastiche d'énoncé du maître Yoda (dont la consonance japonaise du nom – qui, en réalité, vient du sanskrit – est sûrement liée au fait que le japonais place le verbe à la fin de la phrase). L'auteur parle d'*ordre figé ou non de mots*, de *cas*, de *langues SVO/SOV/...* et d'*ambiguïtés syntaxiques* comme l'exemple bien connu « j'ai vu l'homme avec un télescope ». La partie qui concerne la syntaxe est peut-être un tantinet moins élaborée que les autres : après une tentative moyennement réussie d'expliquer en quelques lignes (!) la composition sémantique itérative *bottom-up* d'un arbre syntaxique de constituants, l'auteur se rabat sur la *sémantique lexicale*, les figures (*métaphore, métonymie, synecdoque*) et enfin les *différences de champ lexical* qui rendent difficile la traduction.

Mais, qu'à cela ne tienne, l'auteur se rattrape avec la très riche section consacrée à la pragmatique ! On y apprend que *Babel 17* s'intéresse de près aux *déictiques* et



que dans *Légationville* les phénomènes linguistiques deviennent des référents concrets, des véritables personnages du roman. Ensuite il est question d'*implicatures* (sans les nommer), et enfin, de *performatifs*, illustrés par la nouvelle *L'Histoire de ta vie* de Chiang qui a inspiré le récent film *Premier contact* où les heptapodes connaissent l'avenir et donc tous les actes de leur langage sont performatifs. Puis, on apprend que *Nous autres* de Zamiatine, écrit en Russie en 1920, et ensuite interdit par Staline, décrit une société où les questions sont interdites (idée reprise dans le film *L'Enquête corse*. Enfin, un petit paragraphe sur les *actes locutoires, illocutoires* et *perlocutoires* avec un extrait de *La septième fonction du langage* de Binet.

Pour clore ce chapitre, une section sur la stylistique, illustrée par le bel incipit « *J'avais atteint l'âge de mille kilomètres* » du roman *Le Monde inversé* de Priest, qui a la particularité de se situer dans un monde régi par la géométrie hyperbolique où l'avancée du temps correspond à un mouvement physique. Néanmoins, cette section ne parvient pas à montrer clairement ce que Landragin entend par « stylistique ». Pour ma part, je me serais plutôt attendu à y trouver une mention à la nouvelle *A Two-Timer* (en allemand : *Hausfreund von vorgestern* = L'ami d'avant-hier) de David I. Masson, le récit que fait de notre temps (1968) un voyageur temporel, récit rédigé entièrement en anglais du XVII<sup>e</sup> siècle ! (Notons en passant que David I. Masson, linguiste et accessoirement auteur important de science-fiction, semble avoir totalement échappé à Landragin.) Mais peut-être s'agit-il d'un autre type de stylistique ?

Le cinquième et dernier chapitre de l'ouvrage revient sur son titre : « *Comment parler à un alien ?* » (qui d'ailleurs implique que les aliens sont d'obédience saussurienne et donnent d'office la priorité à la modalité orale de la langue, fait démenti par le film *Premier contact* où l'essentiel de la communication se base sur l'écrit, et pas n'importe quel écrit : un écrit dynamique où la trace écrite évolue dans le temps). Après un récit du décodage des hiéroglyphes égyptiens et quelques exemples afférents de science-fiction (*Expédition* de Boucher, *Omnilingual* de Piper et, plus près de chez nous et de notre adolescence, *La Nuit des temps* de Barjavel), Landragin nous parle de communication à distance avec les aliens et aboutit tout naturellement aux sondes *Voyager* (qui, conformément au phonocentrisme ambiant au moment de leur lancement, contiennent des centaines d'enregistrements audio de diverses langues, mais quasiment aucun texte écrit), leur fameuse plaque représentant une femme et un homme nus, et le film *Contact* où la délicieuse Jodie Foster discute sur une plage exotique mauve avec un alien ayant pris l'apparence de son père défunt.

Suit une section dédiée à la communication entre les espèces, qui comporte plusieurs exemples d'œuvres de science-fiction répertoriant différents types de communication (y compris, dans *Mémoires d'une femme de l'espace* de Mitchison, celle d'une femme qui communique avec un alien dont le corps tout entier, y compris le sexe, participe à la communication, et en tombe enceinte), ainsi que des exemples où la communication n'a simplement pas abouti (comme dans le grandiose *Solaris* de Lem, porté à l'écran par Tarkovski). On y découvre aussi l'existence d'une discipline scientifique appelée « astrolinguistique » (à ne pas confondre avec l'« astroarchéologie » qui est une pseudoscience s'appuyant sur les théories de von

Däniken), discipline dont le but est d'élaborer une langue destinée à la communication avec les aliens, quels qu'ils soient. Après quelques exemples illustrant la difficulté de cette tâche ainsi que celle du face-à-face hypothétique avec les aliens (discipline appelée « xenolinguistique ») et en introduisant les notions de *deixis* et de *multimodalité*, Landragin conclut le chapitre par un historique des « manuels de langue martienne », allant de 1953 jusqu'aux Utopiales de Nantes et à son ouvrage.

Le livre se termine par un épilogue où l'auteur donne des idées d'extension des propriétés des langues existantes et finit par une exhortation aux nouvelles générations d'auteurs d'imaginer des nouvelles de linguistiques-fictions, pour notre plus grand plaisir. Une bibliographie de douze pages et un index des notions concluent cet ouvrage, que je conseille vivement à tout linguiste, ainsi qu'à tout amateur de science-fiction.