

Pour une analyse linguistique des termes dans la phrase

Les vocabulaires scientifiques ou techniques sont consignés dans des nomenclatures multilingues, des dictionnaires spécialisés et des terminologies. L'objet de ces recueils n'est pas de décrire les langues de spécialité dans leur usage mais d'apporter des outils terminologiques fiables aux étudiants et aux experts de chaque secteur d'activité. Un projet ambitieux, celui d'un observatoire qui permettrait de dégager non plus seulement des mots mais un ensemble de collocations grammaticales susceptibles de rendre compte de la grammaire et de la sémantique de ces unités dans la phrase, renouvelle la tradition terminologique jusqu'alors enclavée dans l'étude de sous-systèmes linguistiques.

Termes-clés :
collocations grammaticales ;
sémantique des termes dans la phrase.

Ce bref article qui présente une étude linguistique des termes dans la phrase offre un prolongement au numéro thématique de *Terminologies nouvelles* n° 10 consacré à la phraséologie. Il s'agit, en fait, d'une approche particulière de la phraséologie telle que la conçoit la norme Iso 1087-[Iso 89] dans la définition «Étude scientifique des notions et des termes en usage dans les langues de spécialité». Étude *scientifique*, certes, puisqu'il s'agit d'une approche linguistique, *notionnelle* dans la mesure d'une sous-catégorisation en traits syntactico-sémantiques que nous développerons plus bas, et *en usage* du fait du corpus de référence.

Cette approche linguistique des langues spécialisées aborde les unités par leur fonctionnement au sein des phrases, car la sémantique de ces mots, bien que relevant d'univers particuliers de la connaissance, dépend aussi de l'interprétation des relations grammaticales. Ce qui change de la langue générale dans les langues spécialisées, c'est la phraséologie. Il s'agit donc de proposer une étude qui, partant d'une syntaxe de position, permettra de rendre compte du style statistiquement dominant de tel ou tel corpus. Les descriptions offertes par les opérateurs linguistiques permettent de sélectionner des arguments par domaines d'activité.

Un opérateur peut avoir trois formes : verbe, nom ou adjectif :

- Un verbe dans *Luc est vacciné*;
- Un nom dans *La vaccination de Luc*;

- Un adjectif dans *Luc est vaccinable*.

Un verbe donné est interprété en fonction de son sujet et de son complément, ce qui permet de désambiguïser les phrases :

- *Abaiss*er dans *L'étudiant abaisse le store*; est différent de *abaisser* en géométrie, dans *L'étudiant abaisse une perpendiculaire*;

- *Coder* est transitif dans *Jean code un message*;

- *Coder* est intransitif en biologie, dans *Un gène code pour une protéine*.

L'opérateur est donc défini par l'ensemble de ses arguments.

Les noms prédicatifs sont actualisés par des verbes supports qui ont la même fonction que les désinences verbales.

Dans l'exemple *Le chirurgien procède à l'amputation*, *amputation* est actualisé par *procède*.

Dans *Le navire exécute ses mouvements*, *mouvements* est actualisé par *exécute*.

Un adjectif comme *alcoolique* présente des significations différentes suivant qu'il s'agit :

- *D'une solution alcoolique* «qui par nature contient de l'alcool»;

- *D'une fermentation alcoolique* «particulier à l'alcool éthylique»;

- *D'un délire alcoolique* «qui résulte de l'alcoolisme»;

- *D'un homme alcoolique* «qui s'adonne à l'alcoolisme».

La procédure de traitement (base de données) s'effectue suivant des paramètres ponctuels, préalablement définis suivant chaque sphère d'activité. La réalisation peut se résumer globalement ainsi :

1. Le recensement des vedettes;
2. Le domaine d'emploi;
3. La description des emplois.

1 Le recensement des vedettes

Les ouvrages spécialisés par domaines permettent de recueillir des candidats termes. La présence des unités terminologiques dans les dictionnaires spécialisés confirme l'appartenance des candidats termes à un secteur d'activité donné. Parfois, le recours aux experts s'impose pour des domaines récents qui ne connaissent par de dictionnaires (c'est le cas de la physionique) et pour ceux dont les nombreuses innovations ne permettent pas de disposer de recueils lexicographiques à jour (le secteur de l'informatique par exemple).

Par exemple en partant d'un corpus technique des moyens de transport ferroviaire propres à la Régie autonome des transports parisiens (la RATP), les vedettes et leurs collocations pour les termes *matériel roulant*, MF 77, *matériel ferroviaire* 1977, MS 61, *matériel suburbain* 1961 sont les suivantes :

[1] [...] le MS-61 (*Matériel Suburbain* 1961) est mis en circulation en 1967. [...] il circule exclusivement sur la ligne A [...].

[2] *Traitement des informations à bord du train pour optimiser l'adaptabilité aux différents matériels roulants circulant sur la même ligne.*

[3] *Une nouvelle génération de matériel roulant [...] commençait à circuler sur les lignes régionales [...].*

[4] *Le MF 77, constitué en trains de cinq voitures d'une capacité [...] de 574 voyageurs, circule sur les lignes 7, 8 et 13.*

On a sélectionné pour ces phrases le prédicat verbal *circuler* à partir de leur structure : *circuler*/N0:mt-f/sur N1:loc<voie> où N0 = sujet (matériel de transport ferroviaire) et N1 = locatif (voie de communication).

2 Le domaine d'emploi

La sous-catégorisation en traits syntactico-sémantiques est applicable à l'ensemble des opérateurs décrits au sein d'une classe d'objets. La notion de trait permet de coder sujets et objets spécifiques pour tel opérateur aussi bien en langue générale qu'en langue spécialisée.

[1] *Un pipe-line immergé traversera la Manche pour alimenter les véhicules en carburant.*

[2] *Les trains sont alimentés en courant continu 750 volts [...]*

[3] *Les véhicules étaient désormais alimentés à l'alcool, au gaz de ville ou au gaz de charbon de bois, les fameux «gazogènes».*

1.2.3 = être alimenté/N0:mt-(f+n+tm)/en N1:inc<énergie> où in = «inanimé», c = concret. Les traits syntaxiques sont des outils qui permettent d'apporter des informations complémentaires pour préciser la structure de certains types de phrases et seulement celles-là.

La notion de classe d'objets, développée par le Laboratoire de linguistique informatique (LLI, à l'Université Paris-Nord), permet de prendre en charge le traitement de collocations dans leur milieu contextuel. Elle est chargée d'indiquer le domaine d'emploi (ici mt : moyen de transport) précisé le cas échéant par une rubrique (dans cet exemple : f = transport ferroviaire, n = nautique, tm = terrestre à moteur).

3 La description des emplois

Dans cette étude les phrases sont traitées comme des jeux de prédicats et d'arguments. L'étude syntaxique et sémantique consiste à rendre compte des emplois des unités recensées. Ce type de caractérisation présente des articulations entre syntaxe et

sémantique où le caractère grammatical de l'étude dépend des contraintes de sélection entre un opérateur et son argument nominal.

Le domaine d'emploi du verbe dépend des traits sémantiques révélés par ses arguments.

[1] *L'avion piqua du nez.*

(langage technique);

[2] *Paul piqua du nez.* (langue courante).

Dans la langue générale, les verbes monosémiques sont rares. Dans les langues spécialisées, la sphère d'activité sélectionne le plus souvent un seul emploi, pour un ensemble donné d'arguments, par verbe approprié. Au sein du domaine général des transports on distingue ainsi deux emplois de *descendre* suivant la sphère d'activité (mt = moyen de transport, n = nautique, a = aérien) malgré une structure quasi identique où seulement l'un des deux verbes est transitif. D'autre part la classe d'objets dissocie également les deux emplois de *descendre*.

1. *descendre*/N0:mt-a/sur

N1:loc.<voie>

2. *descendre*/N0:mt-n/N1:loc.<voie>

Pour les contextes :

[1] L'aéronef [...] descendit sur la piste [...].

[2] Ils virent ainsi plusieurs bateaux qui montaient et descendaient péniblement le fleuve.

On peut sélectionner les verbes régissant des noms pertinents pour un domaine d'étude. Si nous choisissons celui des moyens de transport aérien, l'argument *avion* invite certains opérateurs. Pour chaque occurrence de terme, on retrouve soit le verbe, soit le prédicat approprié. Ce sont ces opérateurs appropriés qui définissent les classes d'objets. Il s'agit d'extraire des contextes les prédicats qui fonctionnent et de les coder. Considérons ici une extraction de quelques verbes des moyens de transport :

verbes prédicatifs	corpus	références
atterrir/N0:mt-a/sur N1:loc<voie>	Cet avion [...] est censé décoller et atterrir sur des pistes très courtes.	<i>La Recherche</i> , sept. 94 : 875
circuler/ N0:mt-mtc/sur N2:loc<voie>	C'est fin 1906 [...] que l'omnibus automobile -l'autobus- circula [...] sur la ligne Montmartre/Saint-Germain-des-Prés.	RATP, «Bus - Le matériel roulant» janvier 1990, np
circuler/N0:mt-f/sur N1:loc <voie>	Les premiers trains MS 61 du RER circulaient sur la ligne de Sceaux [...]	RATP, «La RATP a 40 ans», décembre 1988 : 3
circuler/N0:mt-f/sur N1:loc<voie>	Le MF 77, constitué en trains de cinq voitures d'une capacité [...] de 574 voyageurs, circule sur les lignes 7, 8 et 13.	RATP, «La RATP a 40 ans», décembre 1988 : 5
circuler/N0:mt-tmc/sur N1:loc	Les véhicules ont circulé sur l'itinéraire du 94 [...]	RATP, <i>Les transports de surface des origines</i> : 19
circuler/N0:mt-tmc/sur N1:loc<voie>	C'est ainsi que des autobus bleus circulèrent sur la ligne 82 [...]	RATP, <i>Les transports de surface des origines</i> : 18
conduire/N0:hum/N1:mt-tmc	- les machinistes qui conduisent les bus [...]	RATP, «Les années 1990» : 22
être alimenté/N0:mt-f/en N1:inc<énergie>	Les trains sont alimentés en courant continu 750 volts [...]	RATP, «Une journée sur la 1», octobre 1990, np
être alimenté/N0:mt-f/en N1:inc<énergie>/par N2:inc	Le MS 61. Il se présente sous forme d'un élément constitué de 3 voitures : deux motrices, alimentées en courant de traction 1500 volts continu par caténaire [...]	RATP, «RER, le matériel roulant», septembre 1989, np
être alimenté/N0:mt-mt/à N1:inc<énergie>	Les véhicules étaient désormais alimentés à l'alcool, au gaz de ville ou au gaz de charbon de bois, les fameux «gazogènes».	RATP, «De l'omnibus à l'autobus», novembre 1991 : 8
être mis en service/N0:mt-mtc/sur N1:loc<voie>	Juillet 1981 : mise en service sur la ligne 57, d'une voiture prototype [...]	RATP, «Les transports de surface des origines à nos jours», 1985 : 24
être mis en service/N0:mt-tmc/sur N1:loc <voie>/Prép N3:temps	En 1981, fut mis en service sur la ligne 304 un autobus standard [...]	RATP, <i>Les transports de surface des origines</i> : 19
être mis en service/N0:mt-tmc/sur N1:loc<voie>	Novembre 1982 : mise en service sur la ligne 91 du premier autobus articulé dit «superbus», d'une capacité de 127 places.	RATP, «Les transports de surface des origines à nos jours», 1985 : 24
être piloté/N0:mt-tm/par N1:inc<automate>	Étape suivante : concevoir des véhicules susceptibles d'être entièrement pilotés par ordinateur.	<i>La Recherche</i> , nov. 94 : 1104
faire le plein/de N0:mt-tm/de N1:inc<énergie>	Le plein de gasoil du véhicule est fait à la rentrée, l'aspirateur passé.	RATP, «Une journée sur le 38», mai 1989, np

fonctionner/N0:mt-tm/à N1:inc<énergie>	Catalyseur pour pots d'échappement destiné pour véhicules fonctionnant au gaz naturel.	<i>La Recherche</i> , avr. 95 : 373
rouler à/N0:mt-tm/N1:inc<énergie>	un autobus roule à l'hydrogène.	<i>La Recherche</i> , oct. 94 : 984
rouler/N0:mt-tm/à N1:inc<énergie>	[...] plus d'un million de véhicules roulent au gaz naturel	<i>La Recherche</i> , avr. 95 : 372
rouler/N0:mt-tm/entre N1:loc/et N2:loc	[...] un engin à traction autonome faisait son apparition, qu'on appela automobile. En 1905, le Salon de l'Automobile présenta les engins de ce type, qui roulèrent entre la Bourse et le Grand Palais [...]	RATP, «De l'omnibus à l'autobus», novembre 1991 : 4
La suite la plus longue peut être également définie par l'ensemble des arguments des prédicats nominaux. Ces derniers admettent parfois un verbe support (codé W:) et peuvent parfois être à l'origine d'un verbe. manœuvre/N0:hum/N1:mt-n/W:faire <i>faire une manœuvre</i> = <i>manœuvrer</i> (un navire)	stationnement/N0:mt-tm/W:être en être en stationnement = stationner vol/N0:mt-a/W:faire L'avion fait un vol = l'avion vole [...] <i>Le louvoyage rapide du sardinier</i> . Louvoyage est un opérateur nominal à un argument : <i>sardinier</i> [...] <i>les conducteurs chargés de la conduite des métros ou de la</i>	<i>surveillance de leur bonne marche en pilotage automatique</i> . <i>Conduite</i> est un opérateur nominal à deux arguments : <i>conducteur</i> et <i>méto</i> et offre la structure : conduite/N0:hum/N1:mt-f-
termes	noms précatifs	corpus
autobus	circulation/N0:mt-tmc/sur N1:loc<voie>	Pour tenter de résoudre le problème, toujours actuel et pressant, de la circulation des autobus [...]
barque de passage	liaison/N0:mt-n/entre N1:loc/et N2:loc/W:effectuer	Les liaisons entre l'île et les rives du fleuve nécessiteront l'établissement de barques de passage qui constitueront ainsi le premier moyen de transport public par eau sur la Seine.
bateau	manœuvre/N0:hum/N1: mt-n/W:faire	Le redressement brutal du bateau [...] présagea de sa grande stabilité. Les essais en mer permirent de le vérifier après que la manœuvre ait dû être réinventée en s'appuyant sur les souvenirs de quelques marins.
bus	conducteur/N0:hum/N1:mt-tm	Des conducteurs de bus ou de métro aux ingénieurs [...] la liste des métiers exercés par le personnel de la RATP est longue [...]
matériel fer	circulation/N0:mt-f/sur N1:loc <voie>	Première circulation en 1967 sur la ligne 3 du nouveau matériel fer MF 67 [...]

méto	conduite/N0:hum/N1:mt-f	[...] en ligne, les conducteurs chargés de la conduite des métros ou de la surveillance de leur bonne marche en pilotage automatique.
MF 77	consommation/N0:mt-f/N1:inc<énergie>	Mais surtout le MF 77 utilisa pour la première fois un «hacheur de courant» à thyristors (késar) permettant une plus grande souplesse de conduite tout en réduisant la consommation de courant et en autorisant le freinage par récupération [...]
navette	vol/N0:mt-e/W:effectuer	Vingt-huit de ces vols seront effectués par navette.
rame	évacuation/de N0:mt-f/par N1:hum	Les mesures immédiates : elles peuvent consister en une évacuation de rame par ses voyageurs, la rentrée en station d'un train immobilisé sous tunnel [...]
rame	garage/N1:mt-f	À partir de 18h30, ont lieu les premiers garages de rames [...]
sardinier	louvoyage/N0:mt-n	Le louvoyage rapide du sardinier a donc étonné à demi.
train	circulation/card N0:mt-f/Prép N1:temps	À partir de 1983, la véritable interconnexion devenait effective avec la circulation de 8 trains à l'heure de bout en bout sur la ligne B [...]
train	roulement/N0:mt-f/sur N1:loc<voie>	[...] le roulement des trains s'effectue sur des pistes [...]
voilier	embarquement/N0:hum/sur N1:mt-n	Ils sont aujourd'hui intégrés au Port-Musée et participent aux rassemblements de voiliers traditionnels. Le plus ancien a transporté des milliers de personnes [...] et de nombreuses classes de mer et du patrimoine y trouvent un embarquement [...]
Ce projet s'inscrit dans les recherches menées par le Laboratoire de linguistique informatique de Paris XIII qui réalise depuis plusieurs années un dictionnaire électronique de la langue française. Pour ma part, partant du constat que les langues de spécialité utilisent de façon particulière la langue générale comme support nécessaire à la communication, une approche linguistique des langues spécialisées		
abordant les unités dans les phrases pour rendre compte des régularités combinatoires constitue un nouvel axe de recherche. Il se situe au confluent de la terminologie et des méthodes pratiquées en linguistique d'analyse de la phrase par la description de la syntaxe et de la sémantique des termes en contexte. Cette approche permet de dégager non plus seulement des mots, mais un ensemble de collocations		
grammaticales dans les phrases par une étude linguistique des langues de spécialité où la sémantique des termes n'est pas indépendante de leur fonctionnement syntaxique. <i>Jacques Boissy, Centre de terminologie et de néologie, Laboratoire de linguistique informatique, Inal-f-CNRS Ura 1576, Université Paris XIII.</i>		