

La comparaison de la structure interne des groupes nominaux ordinaires et des termes nominaux : les groupes prépositionnels et les groupes adjectivaux

Ágoston NAGY

Dans cet article, nous nous proposons de donner un aperçu sur les différences entre les groupes nominaux (GN par la suite) et les termes nominaux surtout en ce qui concerne les spécificités des compléments prépositionnels (y compris la séquence préposition+nom des noms composés) et celles des ajouts adjectivaux qui apparaissent à l'intérieur d'un GN. La présente recherche s'insère dans une recherche plus large dont le but est d'élaborer une application capable d'extraire automatiquement les termes techniques à partir d'un texte brut qu'elle reçoit en entrée.

Comme cet extracteur automatique se base sur des patrons préalablement définis pour extraire les termes-candidats, il faut précisément définir les structures différentes de termes et démontrer les différences entre les GN ordinaires et les termes. Dans cet article, c'est l'approche classique des termes qui est adoptée, c'est-à-dire un terme est une unité lexicale qui ne dénomme qu'un seul concept auquel il est lié dans le discours spécialisé (Cabré 1999, Wüster 1976). Comme le terme dénomme un seul concept, il constitue une unité lexicale ou une catégorie mineure qui a une forme semblable à un nom composé. Par contre, le GN est une catégorie majeure, une unité syntaxique qui dispose d'une tête obligatoire à laquelle sont liés différents compléments, déterminants et/ou ajouts. Cependant, la frontière entre le GN et les termes n'est pas si tranchée, comme nous montrent les exemples en (1) :

- (1) a. une affaire d'État
- b. une affaire de l'État
- c. la chaise de bois de Marie qui est semblable à un vélo

Parmi ces trois éléments, ce n'est que (1a) qui désigne en même temps un GN et un terme, (1b-c) étant des GN. Le terme peut également disposer d'éléments supplémentaires, par exemple des adjectifs, mais seuls les adjectifs qui désignent une sous-classe de la tête nominale peuvent faire partie des termes comme *filaire* en (2a). Comme (2b) ne désigne pas une sous-classe des réseaux, l'adjectif *grand* ne fait partie du terme.

- (2) a. le réseau filaire
- b. le grand réseau

Dans le but de différencier les termes et les GN, nous présentons d'abord la structure de base des GN en français en mettant un accent particulier sur les groupes prépositionnels (GPrép par la suite) et les groupes adjectivaux (GAdj par la suite) qui se rattachent à la tête nominale parce que ce sont deux des points cruciaux où les GN et les termes nominaux diffèrent. Après avoir présenté ces deux constituants de

base des GN, nous faisons la même analyse avec les termes nominaux.

La structure des GN français

Ce chapitre est basé sur Riegel et al. (2009) et décrit la structure interne des GN du français. La partie obligatoire, la tête du GN est le nom mais un GN peut également contenir des ajouts, des compléments et des spécifieurs. Les spécifieurs du GN sont les déterminants qui précèdent toujours le nom et dont le nombre est limité à deux, ce qui est exemplifié par (3) :

- (3) a. les enfants
- b. deux enfants
- c. les deux enfants
- d. tous les enfants
- e. *tous les deux enfants

Les compléments peuvent être divisés en deux groupes majeurs : les groupes prépositionnels (4a) et les propositions soit infinitives (4b) soit complétives (c'est-à-dire des propositions comportant un verbe fini, comme en 4(c)).

- (4) a. la tarte aux prunes
- b. l'idée de vérifier
- c. le fait qu'il soit parti

La distribution et le nombre des compléments au sein d'un GN sont limités : ils ne peuvent se placer qu'après le nom et, dans la plupart des cas, le nom régit un seul complément.

Les ajouts dans le GN peuvent être des GAdj (5a), des propositions relatives (5b) ou des syntagmes prépositionnels (5c) :

- (5) a. un film intéressant / un intéressant film
- b. la personne que nous avons vue
- c. la personne dans l'armoire

La place et le nombre des ajouts ne sont pas aussi contraints que ceux des compléments. Les GPrép et les propositions ne peuvent que suivre la tête nominale tandis que les adjectifs peuvent être prénominaux ou postnominaux aussi. Les ajouts peuvent être d'un nombre indéfini au sein du GN.

Dans ce qui suit, nous passons en revue les spécificités de ces deux constituants du GN, afin de confronter ces spécificités à celles des GPrép et des GAdj qui font partie des termes nominaux.

La distribution des adjectifs au sein du GN

Comme Cinque (1998) constate, le français, comme la plupart des langues néolatines, est une langue ANA, c'est-à-dire les adjectifs peuvent tant précéder que suivre la tête nominale au sein des GN (6a). Par contre, les langues germaniques, comme l'anglais ou l'allemand, sont des langues AN, c'est-à-dire l'adjectif ne peut

que précéder le nom (6b-c).

- (6) a. le grand bateau rouge
- b. the big red ship
- c. * the big ship red

Cette constatation implique que tous les adjectifs peuvent être préposés ou postposés au nom, ce qui n'est pas valable pour tous les cas. Cependant, la place par défaut de presque tous les adjectifs est la position postnominale, et beaucoup d'adjectifs ne peuvent même pas précéder le nom. Selon Riegel et al. (2009), en moyenne un adjectif sur trois est antéposé au nom mais il peut y avoir d'énormes différences entre les différents types de discours (dans la langue littéraire, un adjectif sur deux est antéposé tandis que ce taux est un sur dix dans le langage scientifique).

Parmi les différents types d'adjectifs, ce sont les adjectifs non-classifiants¹ qui peuvent tant précéder que suivre le nom. La place par défaut de ces adjectifs est réservée après le nom, mais ces adjectifs peuvent être postposés s'ils sont accentués ou s'ils sont, en d'autres termes, focalisés (Laenzlinger 2003).

- (7) a. un livre intéressant
- b. un intéressant livre

Dans quelques cas, il semble y avoir une certaine différence sémantique entre l'adjectif précédant ou suivant le nom. Selon Bouchard (1998), les adjectifs se trouvant dans une position postnominale semblent modifier les composantes sémantiques du nom en tant qu'une entité entière alors que le même adjectif, prénominal, tend à modifier les composantes sémantiques spécifiques intérieures au nom. (8) et (9) montrent des cas typiques où l'adjectif prénominal n'équivaut pas à l'adjectif postnominal :

- (8) a. une voiture nouvelle
- b. une nouvelle voiture
- (9) a. un homme seul
- b. un seul homme

Ces exemples nous montrent clairement la différence dans l'interprétation sémantique entre les variantes pré- ou postnominales de ces types d'adjectifs. Par exemple (9a) se réfère à un homme qui est seul – dans sa totalité –, et (9b) désigne un homme qui est seul dans un environnement spécifique, par exemple *Il n'y a qu'un seul homme dans ce groupe*.

Quelques adjectifs précèdent toujours le nom : ce sont en général des adjectifs monosyllabiques ou courts. Laenzlinger (2003) explique ce fait par des facteurs d'utilisation et phonorhythmiques ; en effet, ces noms s'utilisent souvent dans la conversation quotidienne. Ces adjectifs sont *petit*, *beau* ou *long*.

¹ Les adjectifs non-classifiants (par exemple *merveilleux*, *gentil*) désignent des propriétés subjectives et peuvent donc se combiner avec des adverbes d'intensité comme *très*, *trop*.

- (10) a. une petite pause
- b. une belle maison
- c. une petite belle fille

Le GN en (10c) justifie également que même plusieurs de ces adjectifs peuvent précéder le nom. Cependant, si ces adjectifs régissent un complément prépositionnel, ils suivent obligatoirement le nom (11).

- (11) a. une _{AP}[haute] tour
- b. une tour _{AP}[haute de 200 mètres]
- c. une tour _{AP}[plus haute que la Tour Eiffel]
- d. *une _{AP}[haute de 200 mètres] tour

Un adjectif prénominal devient également postnominal s'il est modifié par un adverbe sauf si l'adjectif et l'adverbe sont courts et fréquents (12). Les adverbes de ce type sont entre autres *tout*, *très* ou *trop* dans le cas desquels la position du GAdj est facultative à l'intérieur du GN (12).

- (12) a. un _{AP}[long] voyage
- b. un _{AP}[très long] voyage / un voyage _{AP}[très long]
- (13) a. un voyage _{AP}[extrêmement long]
- b. *un _{AP}[extrêmement long] voyage

Les autres types d'adjectifs, notamment les adjectifs relationnels et les adjectifs classifiants sont toujours postposés au nom. Les premiers sont des adjectifs dérivés d'un nom et paraphrasables par un complément prépositionnel du nom, comme *municipal* ou *solaire*. Les adjectifs classifiants ou les adjectifs prédicatifs intersectifs suivent également le nom (14). Ces derniers désignent des propriétés qui classifient les noms à partir de leur origine, forme, couleur ou appartenance à une communauté. En outre, les adjectifs dérivés d'un verbe, notamment les participes présent ou passé, doivent également suivre le nom (15). (Bouchard 1998)

- (14) a. un bureau ovale
- b. une assiette chinoise
- c. une expérimentation biochimique
- (15) a. un voyage organisé
- b. une chaise pliante

Apparemment, les adjectifs intensionnels dérivés contredisent la règle selon laquelle la position postnominale est la place par défaut des adjectifs parce que ceux-ci ne peuvent être que prénominaux en français, ce qui est exemplifié en (16).

- (16) a. un prétendu chef-cuisinier
- b. un soi-disant médecin

Il serait intéressant d'analyser la distribution des ajouts adjectivaux et des GPrép au sein du GN pour pouvoir énumérer exhaustivement tous les patrons possibles des

termes nominaux. La place du GPrép n'est pas problématique dans la mesure où il suit toujours la tête nominale. En revanche, la position respective du GPrép et du GAdj n'est pas toujours facile à déterminer.

Selon Laenzlinger (2003), un adjectif peut s'insérer entre la tête nominale et le GPrép complément mais il peut également se placer après le GPrép complément, sa place est donc facultative, ce qui est exemplifié en (17) :

- (17) a. un ministre de la défense noir
- b. un ministre noir de la défense

Cependant, si la tête nominale et le GPrép complément désignent ensemble une entité lexicalement figée, l'adjectif ne peut pas s'insérer entre les constituants de cette unité, ce qui est montré en (18) où *pomme de terre* forme une unité lexicale.

- (18) a. des pommes de terre nouvelles
- b. ??des pommes nouvelles de terre

Si le GPrép est un ajout à l'intérieur du GN, le GAdj se situe avant celui-là (19a), même s'il est complexe (20a), sinon il modifierait le GN se trouvant à l'intérieur du GPrép ajout (19c et 20c).

- (19) a. une maison _{GAdj}[blanche] _{GPrép}[sur la colline]
- b. *une maison _{GPrép}[sur la colline] _{GAdj}[blanche]
- c. une maison _{GPrép}[sur _{GN}[la colline _{GAdj}[blanche]]]
- (20) a. une maison _{GAdj}[haute de 100 mètres] _{GPrép}[sur la colline]
- b. *une maison _{GPrép}[sur la colline] _{GAdj}[haute de 100 mètres]
- c. une maison _{GPrép}[sur _{GN}[la colline _{GAdj}[haute de 100 mètres]]]

Abeillé et Godard (1999) analysent d'un autre point de vue la distribution relative des têtes nominales et leurs ajouts adjectivaux. Elles introduisent la notion de *poids relatif* qui leur permet d'expliquer la place des adjectifs par rapport à la tête nominale. Elles proposent deux traits pour classer les adjectifs suivant leur poids : *lite* 'léger' ou *non-lite* 'lourd'. Le poids relatif d'un adjectif peut être une information stockée dans le lexique (certains adjectifs sont légers, d'autres non, et pour de nombreux adjectifs cette propriété n'est pas déterminée), mais il peut aussi être déterminé par des règles qui décrivent la structure syntaxique des syntagmes (la plupart des syntagmes sont lourds, d'autres légers). Il existe notamment une règle qui veut que les adjectifs légers soient obligatoirement prénominaux, les lourds postnominaux.

- (21) a. une _{lite} A[jolie] dame
- b. une femme _{non-lite} A[chinoise]

Cependant les adjectifs légers ne peuvent pas modifier les expressions nominales lourdes (comme un GN formé par deux ou plusieurs GN coordonnés) : dans ce cas-là ils ne peuvent être que lourds, c'est-à-dire ils ne peuvent que suivre la tête nominale, ce qui est prouvé par (22) :

- (22) non-lite N[des femmes et des hommes] non-lite A[jolis]

Le poids relatif des adjectifs coordonnés n'est pas toujours univoque. Le poids de deux adjectifs légers coordonnés n'est pas déterminé, ils peuvent donc et précéder et suivre le nom, ce qui est démontré en (23), et la coordination de deux adjectifs dont le poids n'est pas déterminé par le lexique correspond à une unité lourde, ce qui est exemplifié en (24).

- (23) a. une A non.dét.[lite-A[belle] et lite-A[jolie]] dame
b. une dame A non.dét.[lite-A[belle] et lite-A[jolie]]

- (24) a. ??une non-lite A[A non.dét.[magnifique] et A non.dét.[exceptionnelle]] dame
b. une dame non-lite A[A non.dét.[magnifique] et A non.dét.[exceptionnelle]]

Les compléments prépositionnels du GN

La définition du GPrép n'est pas simple à formuler comme il n'est pas toujours facile de distinguer un GPrép introduisant une nouvelle entité à l'intérieur du GN (25a) et celui qui est associé à la tête nominale pour former avec celle-ci un nom composé (25b). Dans le cas des noms composés (25b), la préposition est généralement suivie d'un nom sans déterminant parce que la présence du déterminant impliquerait un GN complexe où le GN précédé d'une préposition peut être considéré comme un GN enchâssé qui a une référence à part (25a). Cependant, les GPrép dans les noms composés ne sont pas référentiels. Le déterminant a donc ici un rôle essentiel dans la différenciation des noms composés de type *N Prép N* et les GN comportant un GPrép.

- (25) a. le moulin du village
b. le moulin à vent

Riegel et al. (2009) rangent parmi les GPrép tout groupe qui se compose d'une préposition suivie d'un groupe nominal entier (comme *le chien GPrép[de GN[la voisine]]*). Cependant, parmi leurs exemples, on peut trouver des noms composés comme *canne à pêche* où *pêche* ne constitue pas seul un GN. Pourtant, il existe des points de vue (par exemple Bosredon et Tamba 1991) selon lesquels les noms composés sont des noms simples sur le plan sémantique mais ils constituent un GN sur le plan formel. De cette façon, Bosredon et Tamba (1991) distinguent les GPrép des séquences préposition+nom reliés à un nom en les nommant constituants et formants respectivement.

Par la suite, nous entendons par GPrép et les constituants et les formants prépositionnels parce que la frontière entre formant et constituant n'est pas nette du point de vue formel. D'une part, il existe des prépositions qui ne sont pas suivies de déterminant en général (par exemple *par* ou *en*), et ne forment pas avec le nom précédé d'un nom composé (26a). D'autre part, il existe des noms composés qui contiennent un GPrép où la préposition est suivie d'un déterminant (27a).

- (26) a. voyage en Italie
b. *voyage en l'Italie

- (27) a. cancer de la peau
b. ?cancer de peau

Par la suite, nous nous concentrons sur les noms composés sans déterminant, c'est-à-dire ceux qui sont plus aptes à devenir des termes.

Dans le français, les noms composés sont réalisés en général par des noms (28a-d) ou infinitifs (28e) reliés à la tête nominale par l'intermédiaire d'une préposition. Les mots composés se forment généralement avec l'utilisation de la préposition *de* (28a-b), mais dans certains cas, les deux éléments peuvent être reliés par d'autres prépositions (comme *en* dans (28c) ou *à* dans (28d-e)). Les mots composés se forment en général sans trait d'union sauf quelques exceptions comme (28c). (Riegel et al. 2009)

- (28) a. hôtel de ville
b. professeur de français
c. arc-en-ciel
d. rouleau à pâtisserie
e. machine à laver

L'importance de la présence du trait d'union n'est pas seulement une question d'orthographe. L'extracteur terminologique que nous élaborons s'appuie sur des annotations automatiques, et les annotateurs (y compris celui qui est utilisé dans cette application) ne séparent pas les éléments reliés par des traits d'unions mais les considèrent comme un mot appartenant à la catégorie du *nom*. De cette façon, les mots composés de type (28c) s'inscrivent dans le même patron syntaxique que celui qui s'applique aux termes se composant d'un seul nom comme *réseau*.

Le français connaît également les mots composés sans préposition qui peuvent s'écrire avec (29a-b) ou sans traits d'union (29c).

- (29) a. le gratte-ciel
b. le chou-fleur
c. la pause déjeuner

Les GPrép dans les termes nominaux

Parmi les différents compléments ou ajouts dont un GN est susceptible de disposer, c'est le GPrép et le GAdj qui peuvent apparaître à l'intérieur des termes nominaux. Les propositions à verbe fini ne font pas partie des termes car ils introduisent au moins une nouvelle entité. Par exemple dans (30), dans la proposition relative apparaît un nouveau concept, *l'utilisateur*.

- (30) a. le site web que l'utilisateur a visité

Un terme nominal ne contient pas donc une proposition relative mais il peut bien sûr être paraphrasé par une proposition relative, comme nous montre (31a-b), mais la

version susceptible de devenir terme est sans aucun doute (31a).

- (31) a. réseau filaire
- b. un réseau qui est filaire

Parmi les différents compléments, les termes ne peuvent avoir qu'un GPrép comme complément, c'est pourquoi c'est uniquement ce type de complément qui est analysé dans cet article. Comme nous l'avons décrit plus haut, les noms composés prépositionnels lexicalisés en français en général ne contiennent pas de déterminants internes au sein du GPrép complément et la préposition est suivie d'un nom (le terme en 32a) ou d'un infinitif (le terme en 32b).

- (32) a. farine de blé
- b. machine à laver

Ces éléments composés doivent être considérés comme une unité lexicale puisque *farine* et *blé* sont des unités lexicales qui peuvent apparaître comme des unités autonomes mais s'ils se trouvent reliés par la préposition *de*, ils désignent un troisième concept à part. Cependant, si la préposition est suivie d'un déterminant, la situation deviendra plus complexe. Par défaut, un GPrép complément comportant un déterminant ne doit pas être analysé comme un élément constitutif du terme nominal parce que dans ce cas-là, le GN inclus dans le GPrép désigne une entité à part, comme en (33) :

- (33) a. mise à jour du site web
- b. création d'un site web

En ce cas-là, il est plus raisonnable d'analyser ce GN comme comportant deux termes indépendants, notamment *mise à jour* et *site web*. Les déterminants fonctionnent donc comme des séparateurs entre les différents termes.

Cadiot (1993) soutient également ce point de vue : un GPrép sans déterminant désigne une sous-classe du nom précédent alors que le GPrép contenant un déterminant ne décrit qu'une occurrence du nom qui le précède. Dans le dernier cas, la classification ne peut être qu'indirecte : cela découle de la propriété extensionnelle du GPrép à déterminant, tandis que dans le premier cas, le GPrép peut classer le nom sur la base de ses propriétés intensionnelles.

- (34) a. arbre à feuilles caduques
- b. arbre aux feuilles jaunies

Les exemples en (34) montrent que le (34a) sans déterminant désigne une sous-classe des arbres, tandis que la version à déterminant (34b) décrit un arbre qui a des feuilles jaunies mais qui ne constitue pas une sous-classe des arbres.

Anscombe (1990, 1991) adopte le même point de vue lorsqu'il constate que le GPrép sans déterminant décrit une propriété essentielle de la tête nominale, tandis que le GPrép à déterminant en décrit une propriété accidentelle. Selon lui, une propriété P est une propriété essentielle de l'entité E si P peut être considérée comme une unité inaliénable de E. Par contre, P est une propriété accidentelle s'il

est une propriété temporaire. La propriété essentielle est donc une propriété innée, tandis que la propriété accidentelle n'est qu'un état actuel. Les exemples en (35) montrent que dans le cas de *bateau à voiles*, qui contiennent un GPrép sans déterminant, seuls quelques adjectifs peuvent être utilisés pour modifier le nom sans déterminant qui suit la préposition : ces adjectifs doivent désigner le type de la voile (35b) tandis que dans la version avec déterminant (35d), ceux-ci doivent désigner un état actuel de la voile.

- (35) a. bateau à voiles
- b. bateau à voiles carrés/latines
- c. bateau à voiles ??hissés/*déchirés
- d. bateau aux voiles hissés/déchirés

Cadiot (1993) observe que la situation est similaire si le GPrép ne contient pas d'adjectifs. La présence du déterminant implique que l'élément précédant et suivant la préposition désigne une entité pourvue d'une référence autonome. Ceci est justifié par les paires suivantes :

- (36) a. un sac à dos
- b. un sac au dos
- (37) a. Paul a un sac à dos mais il le porte à la main.
- b. *Paul a un sac au dos mais il le porte à la main.

Anscombe (1991) constate que le GPrép sans déterminant ne peut pas comporter n'importe quel nom mais un nom qui désigne une propriété non évidente de la tête. Par exemple, une vente est intrinsèquement à bénéfice, mais si cette opération aboutit à une perte, ceci est une propriété non-évidente.

- (38) a. vente à *bénéfice
- b. vente à perte
- (39) a. *un mouton à quatre pattes
- b. *une voiture à moteur

Les exemples en (39a-b) sont incorrects pour les mêmes raisons : les moutons ont quatre pattes, les voitures ont un moteur – ce sont des propriétés essentielles. Si ces compléments sont modifiés ou élargis, nous obtiendrons des GN corrects car les GPrép de *un mouton à cinq pattes* ou *un mouton à moteur à alcool* désignent des propriétés non évidentes de la tête nominale.

- (40) a. un mouton à cinq pattes
- b. une voiture à moteur à alcool

Toutefois, la constatation de Cadiot (1993) selon laquelle le déterminant a pour fonction d'introduire une nouvelle entité n'est pas toujours vraie. En effet, il existe des termes qui contiennent un déterminant devant la composante nominale qui n'exprime pas une entité à part, ce qui est exemplifié en (41) :

- (41) a. cancer de la peau
b. vidéo à la demande

Il serait difficile d'expliquer pourquoi *cancer de la peau* au lieu de *cancer de peau*. Dans une étude antérieure (Nagy 2009), nous avons démontré que le taux du GN à déterminant interne est d'environ 7% par rapport à la totalité des termes nominaux, mais nous n'avons pas calculé le taux des GN à déterminant qui peuvent apparaître également sans déterminant. Par conséquent, tout en étant conscient de la perte que cela représente, nous n'allons pas considérer les termes avec déterminant lors du processus de l'extraction automatique.

Une autre caractéristique intéressante des GPrép dans les termes est que la préposition peut être absente (par exemple en (42)). Cette observation est surtout valable pour les cas des termes récemment créés dont Béjoint et Ahronian (2008) constatent que l'omission est due à l'effet de l'anglais. Toutefois, l'ordre des noms est plutôt conforme aux règles du français.

- (42) a. code source
b. accès Internet

Ceci ne constitue pas un problème pour notre extracteur automatique, car le patron reconnaissant les noms composés va les reconnaître sans modification, du fait que le schéma N+N caractérise également quelques termes.

Les GAdj dans les termes nominaux

La place des adjectifs est un point crucial quand il s'agit de décider s'ils apparaissent dans un terme ou non. Comme nous l'avons évoqué précédemment, la place par défaut des adjectifs dans un GN est la position postnominale, mais certains adjectifs peuvent apparaître dans une position prénominale s'il s'agit d'une mise en relief par exemple. L'accentuation liée à des raisons émotives ne joue aucun rôle dans la langue de spécialité, car celle-ci exige une forte objectivité : elle utilise avant tout des adjectifs classifiants et relationnels et ne recourt pas à l'accentuation d'insistance. Ceci est exemplifié en (43), où l'adjectif relationnel ne peut pas être antéposé au nom à cause des caractéristiques des adjectifs relationnels :

- (43) a. un réseau filaire
b. *un filaire réseau

Les adjectifs monosyllabiques fréquemment utilisés ont également peu de chances d'apparaître dans un terme parce qu'ils désignent plutôt une caractéristique accidentelle des termes. C'est aussi le cas pour les adjectifs intensionnels dérivés d'un verbe (par exemple *soi-disant*), qui peuvent également précéder un terme mais ils n'en font jamais partie.

- (44) a. un grand réseau filaire
b. un prétendu réseau filaire

Dans une étude antérieure (Nagy 2009), nous avons constaté qu'il n'y avait pas de

terme qui commence par un adjectif préposé dans le corpus informatique dont nous nous sommes servi. Nous excluons en effet cette possibilité même si d'autres termes en d'autres domaines spécialisés peuvent contenir des adjectifs préposés :

- (45) a. petite aiguille
- b. premier ministre

Dans (45b) l'adjectif ordinal précède le nom, comme la plupart des adjectifs ordinaux, mais ce type d'adjectif a également un rôle d'organiseur de texte ou anaphorique, c'est-à-dire il reprend une occurrence spécifique d'un nom déjà mentionné et donc ne fait pas partie du terme.

Hormis nos études, nous n'avons pas de données mesurées sur la proportion des termes commençant par un adjectif, c'est pourquoi, sur la base de l'étude susmentionnée, cette possibilité ne sera pas prise en compte.

Conclusion

Dans cet article, nous avons revu la structure interne des GN en les comparant avec les termes nominaux concernant leurs ajouts et compléments possibles, en nous concentrant sur les GAdj et les GPrép. Cette comparaison a été faite en vue de l'établissement d'une liste exhaustive des différentes structures syntaxiques possibles des termes nominaux pour que notre extracteur automatique des termes reconnaisse le plus de termes possibles.

Les termes nominaux n'acceptent pas en général les GAdj préposés car la langue spécialisée n'admet que les adjectifs classifiants et relationnels qui ne peuvent être que postposés aux noms. Dans quelques cas, les termes peuvent admettre des adjectifs préposés, notamment les adjectifs monosyllabiques comme *petit* ou des adjectifs ordinaux mais si ces types d'adjectifs précèdent le nom, dans la plupart des cas, ils ne désignent pas une qualité existentielle du nom et par conséquent ne peuvent pas faire partie des termes nominaux.

Les compléments des termes nominaux en général ne peuvent pas être des GPrép à déterminant puisque le déterminant introduit une nouvelle entité, un nouveau concept et les termes ne peuvent désigner qu'un seul concept. Même si quelques termes peuvent comporter des GPrép comportant un déterminant, leur taux est si infime qu'il ne serait pas raisonnable de les prendre en considération car cela produirait beaucoup de bruit durant l'extraction automatique.

Références bibliographiques

- ABEILLE, A., GODARD, D., 1999, « La position de l'adjectif en français : le poids des mots », *Recherches Linguistiques de Vincennes* (28), 9–32.
- ANSCOMBRE, M. J.-C., 1990, « Article zéro et structuration d'événements », in *Le discours : représentations et interprétations*, sous la dir. de M. Charolles & J. Jayez, PUN, 265–300.
- ANSCOMBRE, M. J.-C., 1991, « L'article zéro sous préposition », *Langue française* (91), 24–39.
- BEJOINT, H., AHRONIAN, C., 2008, « Les noms composés anglais et français du domaine d'Internet : une radiographie bilingue », *Meta : journal des traducteurs*, (53/3), 648–666.
- BOSREDON B., TAMBA I., 1991. « Verre à pied, moule à gaufres : préposition et noms composés de sous-classe », *Langue française* (91), 40–55.
- BOUCHARD, D., 1998, « The distribution and interpretation of adjectives in French: a consequence of Base Phrase Structure », *Probus* (10).
- CABRE, M. T., 1999. *Terminology. Theory, methods and applications*, Amsterdam / Philadelphia, John Benjamins.
- CADIOT, P., 1993, « *À* entre deux noms: vers la composition nominale », *Lexique* (11), 193–240.
- CINQUE, G., 1994, « On the evidence for partial N-movement in the Romance DP », in *Path Towards Universal Grammar. Studies in Honor of Richard Kayne*, sous la dir. de Cinque, G. Washington, Georgetown University Press, 85–110.
- LAENZLINGER, C., 2003, *Initiation à la Syntaxe formelle du français : Le modèle. Principes et Paramètres de la Grammaire Générative Transformationnelle*, Berne, Peter Lang AG.
- NAGY, Á., 2009, « La structure interne des termes techniques du français et leur reconnaissance par ordinateur », in *La perspective interdisciplinaire des études françaises et francophones*, sous la dir. de Anna Kieliszczyk, Ewa Pilecka, Łask, Oficyna Wydawnicza LEKSEM, 117–123.
- RIEGEL, M., PELLAT, J.-Ch., RIOUL, R., 2009, *Grammaire méthodique du français* (4^e édition), Paris, PUF.
- WÜSTER, E. 1976. « La théorie générale de la terminologie, un domaine interdisciplinaire impliquant la linguistique, la logique, l'ontologie, l'informatique et la science des objets », *Actes du colloque international de terminologie, Québec 5–8 octobre 1975*, Québec, L'Éditeur officielle du Québec.