

À propos de la traduction automatique des déterminants de l'espagnol et du français

XAVIER BLANCO

Université Autonome de Barcelone, Barcelone, Espagne

PIERRE-ANDRÉ BUVET

Université de Franche-Comté, Besançon, France

RÉSUMÉ

Dans un premier temps, nous discutons des principales particularités des contraintes sur les déterminants du français et de la façon de les représenter formellement en vue notamment de leur intégration dans un système de traduction automatique franco-espagnol. Nous examinons ensuite les principes de fonctionnement des modules de correspondance entre les deux langues qu'il convient d'élaborer.

ABSTRACT

The author's first discuss major constraints regarding determination in French and how they might be formalized for integration into a French-Spanish automatic translation system. They then looks at the guiding principles for developing the necessary correspondence modules.

Le projet DétTAL (Détermination et Traitement Automatique du Langage) s'est fixé comme objectif de développer des modélisations de la détermination de l'espagnol et du français ainsi que des modules de correspondance, relatifs aux déterminants, entre les deux langues en vue de leur implémentation dans des systèmes de traduction automatique¹. Nous présentons tout d'abord les particularités théoriques et méthodologiques du projet DétTAL. Nous discutons ensuite des particularités des modules de traduction qui rendent compte des contraintes sur les déterminants.

1. LE PROJET DétTAL : REMARQUES THÉORIQUES ET MÉTHODOLOGIQUES

À l'origine de DétTAL, il y a un constat : le domaine du TALN (Traitement Automatique des Langues Naturelles) comporte peu de travaux sur la détermination alors que le traitement des déterminants est essentiel pour la traduction et la génération automatique. Un tel manque ne serait pas dû à la simplicité des modélisations à effectuer mais plutôt à leur complexité².

À la suite de ce constat, nous avons cherché à rendre compte de la détermination dans une perspective contrastive, notre objectif final étant de développer des applications pour la traduction automatique. Dans cette optique, l'étude de la détermination est réductible à celle des contraintes sur les déterminants d'un nom donné. L'analyse de ces contraintes comporte deux volets :

- la formalisation des contraintes sur les déterminants de l'espagnol et du français dans le cadre général d'une grammaire comparée de ces deux langues ;
- l'élaboration formelle des correspondances entre les contraintes décrites respectivement pour l'espagnol et le français afin qu'un système de traduction automatique puisse les prendre en compte.

Dans cette section, nous ne considérons que le premier volet. Nous discutons en premier lieu de la spécificité de notre analyse de la détermination; nous proposons d'étudier les déterminants en tenant compte de la bipartition syntaxique des substantifs résultant de la décomposition d'une phrase simple en un prédicat et ses arguments. Nous proposons ensuite de recourir systématiquement à la notion de classes d'objets pour décrire les contraintes sur les déterminants. Pour l'instant, les exemples que nous donnons portent uniquement sur le français. Les déterminants de l'espagnol seront traités dans la seconde partie.

1.1. Une analyse « atypique » de la détermination

Après une rapide présentation des fondements théoriques de nos travaux, nous indiquons les données qui justifient le traitement de la détermination que nous proposons ensuite.

1.1.1 *Fondements théoriques*

Nous n'avons pas retenu les modèles logico-référentiels que plusieurs linguistes ont développés³. Dans les études qui sous-tendent ces modèles, on fait peu de cas de la diversité des substantifs pour expliquer leur compatibilité avec les déterminants car, d'une manière générale, la dimension lexicale de la détermination n'est pas prise en compte. De plus, le cadre d'analyse des déterminants est souvent restreint au seul groupe nominal; un tel choix, assurément justifié dans une perspective référentielle, ne permet pas de rendre compte des nombreuses contraintes sur les déterminants qui résultent des autres éléments de la phrase. Bien que nous réfutons ces modèles, nous reconnaissons que certaines de leurs représentations conservent une grande pertinence⁴.

Le modèle morpho-syntaxique développé par Maurice Gross dans le cadre du lexique-grammaire⁵ a constitué le point de départ de nos travaux. Ce modèle s'attache à décrire la diversité des combinaisons entre les noms et leurs déterminants en les analysant dans le cadre de la phrase. Toutefois, nous estimons que le recensement des déterminants⁶ et des diverses transformations qui leur sont attachées est insuffisant dans une perspective comparatiste; il faut prendre en compte également la dimension syntactico-sémantique de la détermination. C'est pourquoi, afin de pouvoir traduire automatiquement dans une langue L_j la détermination d'un nom donné dans une phrase donnée d'une langue L_i, nous avons développé notre propre modèle où les différentes formes de déterminants⁷ sont associées à des catégories qui donnent lieu à des interprétations sémantiques.

Le tableau ci-dessous rend compte de différentes situations relatives à la détermination du français; les catégories proposées doivent être considérées comme une simple base de travail et non une typologie définitive, qui reste à faire.

La classification que nous proposons suggère une approche uniforme des déterminants. Nous allons montrer que cette homogénéité n'est qu'apparente et que notre analyse des déterminants est moins usuelle qu'il n'y paraît. Notons par ailleurs que toutes ces catégories ne sont pas sur le même plan. Nous distinguons en premier lieu la catégorie des définis de toutes les autres catégories, que nous considérons avant tout comme se rapportant à des déterminants non définis⁸.

Catégorie de déterminants	exemples en français
l'article zéro	<i>zéro, zéro-Modif</i>
le partitif	<i>du, des</i>
les comparatifs	<i>autant (de) ... que, plus (de) ...que, ...</i>
les définis	<i>ce, le, le-Modif, Poss, ...</i>
les génériques	<i>le, les, un</i>
les indéfinis	<i>un, un-Modif, ...</i>
les intensifs	<i>beaucoup de, très, un tas de, un de ces, ...</i>
les quantifieurs	<i>beaucoup de, plusieurs, trois, un tas de, ...</i>

Nous devons également au lexique-grammaire la représentation de la phrase simple que nous avons adoptée; un phrase simple équivaut à un opérateur, ou prédicat, et son domaine d'arguments. Considérons l'énoncé suivant :

- (1) *Une fille a giflé un garçon*

Cette construction phrastique est associée à l'opérateur *gifler* et le domaine d'arguments binaires qui le caractérise, la paire (*fille, garçon*) est une des solutions plausibles en termes d'arguments. Cette interprétation de la phrase simple et les notions de verbe support et de nom prédicatif qui s'en sont suivies⁹ ont permis, entre autres, de renouveler les travaux sur les substantifs en brisant le carcan des catégories grammaticales et de ne plus concevoir les verbes qu'en position dominante dans la phrase. Ainsi, si l'on admet que (1) est synonyme de :

- (2) *Un fille a donné une gifle à un garçon*

il apparaît clairement que l'opérateur ne se présente plus sous la forme d'un verbe mais sous celle d'un nom; *gifle* s'interprète ici comme un prédicat nominal actualisé par le verbe support *donner*. Ces analyses sont bien connues, aussi contentons-nous de rappeler que les noms prédicatifs ne sont pas nécessairement rattachés morphologiquement à des verbes ou à des adjectifs (voire les deux); c'est le cas par exemple de *leçon* dans :

- (3) *Luc a donné une leçon à Max*

Une des conséquences majeures de ces nouveaux travaux a été la subdivision des noms en deux classes syntaxiques disjointes: celle des arguments élémentaires, d'une part, celle des prédicats, d'autre part.

1.1.2 Les données linguistiques

Les études systématiques portant sur les noms prédicatifs¹⁰ ont montré qu'il existait, dans le cadre des constructions à support, de nombreuses contraintes relatives à la détermination de ces substantifs. Ainsi, pour ne prendre qu'un exemple, celui des noms de maladies précédés du verbe *avoir*, on constate que les substantifs *grippe*, *hépatite* ou *variole* n'ont pas une détermination homogène :

- (4) *Luc avait (la + une) grippe*

- (5) *Luc avait (*la + une) hépatite*

- (6) *Luc avait (la + *une) variole¹¹*

Du fait que les contraintes de ce type sont spécifiques aux seuls prédicats nominaux (du moins à certains d'entre eux), nous avons postulé un traitement différencié de leurs déterminants.

La nécessité d'une bipartition de la détermination apparaît également au regard d'une particularité des seuls déterminants des noms prédicatifs ; certains d'entre eux peuvent caractériser aspectuellement les opérateurs nominaux qui les acceptent. Ainsi, dans l'énoncé suivant :

(7) *Luc a fait plusieurs repas*

le déterminant *plusieurs* appliqué à *repas* fait état de plus d'une occurrence de l'action signifiée comme l'indique la paraphrase suivante :

(7') *Luc a fait plusieurs fois un repas*

De ce fait, on l'interprète comme un marqueur aspectuel¹². Les correspondances entre des adverbes (simples ou complexes) et certains déterminants, relatifs à un même prédicat, sont des phénomènes qui ont été identifiés sous le nom de « descente de l'adverbe¹³ », comme dans le cas des phrases ci-dessous que l'on considère comme équivalentes :

(8) *Luc a fait un achat de livres plusieurs fois*

(8') *Luc a fait plusieurs achats de livres*

Du point de vue aspectuel, le déterminant *plusieurs* de (8') et l'adverbe *plusieurs fois* de (8) rapportés à *achat* ont la même valeur : le fréquentatif.

Au même titre qu'un adverbe ou un déterminant, un verbe support peut véhiculer une information aspectuelle relative au prédicat qu'il actualise. Dans de tels cas, les déterminants du substantif prédicatif peuvent présenter des contraintes spécifiques¹⁴ :

(9) *Luc a accumulé (les + *des) erreurs*

Le verbe *accumuler* est considéré ici comme une variante, dite aspectuelle, du verbe support standard *faire* :

(9') *Luc a fait (*les + des) erreurs*

Différents moyens grammaticaux permettent donc de caractériser aspectuellement un prédicat. Harris (1976), dont les travaux ont inspiré ceux du lexique-grammaire, considère que l'aspect est un opérateur de niveau supérieur (du type Oo), c'est-à-dire un opérateur qui s'applique à une phrase. Le fait que les déterminants, pour ce qui est des substantifs prédicatifs, soient l'un de ces moyens grammaticaux contribue à justifier le traitement différencié de la détermination que nous proposons.

1.2. Modélisation des contraintes sur les déterminants

Nous allons montrer à présent l'intérêt d'envisager deux sous-systèmes de déterminants du point de vue de la représentation de la phrase proposée dans le lexique-grammaire. La décomposition syntactico-sémantique de la phrase simple en un prédicat et son domaine d'arguments donne lieu à la formulation suivante :

$P \rightarrow \text{Préd} (\text{Arg.})^{15}$

Si l'on applique ce type de schématisation aux phrases suivantes :

(10) *Un enfant a mangé plusieurs pommes*

(11) *Un enfant a fait plusieurs bêtises*

on obtient respectivement :

$P_{(10)} \rightarrow$ manger (enfant, pomme)

$P_{(11)} \rightarrow$ bêtise (enfant)¹⁶

Une telle représentation formelle de la phrase simple permet de rendre compte d'une grande variété de faits de langue et, de plus, elle permet d'effectuer automatiquement diverses opérations portant sur des données linguistiques¹⁷.

Jusqu'à présent, nous avons considéré le moyen de représenter formellement un énoncé phrastique. Désormais, nous examinons la situation inverse, c'est-à-dire la possibilité de produire des phrases grammaticales à partir d'une phrase schématisée. Généralement, deux opérations, la linéarisation et l'actualisation, sont proposées. La première consiste à établir la structure syntaxique de la phrase schématisée en ordonnant les différents arguments par rapport au prédicat. Pour nous, l'actualisation est un simple procédé grammatical (une mise en phrase) et non un procédé référentiel. Cette opération consiste à :

- conjuguer les prédicats verbaux et, par le biais des verbes supports, les prédicats adjectivaux ou nominaux ;
- (s'il y a lieu) préciser les mots grammaticaux, c'est-à-dire les déterminants et les prépositions éventuelles qui indiquent le mode de rection entre les prédicats et certains de leurs arguments.

Cette dernière façon de procéder ne nous paraît pas totalement satisfaisante, car elle tend à uniformiser des phénomènes qui ne sont pas de même nature. Ainsi, pour ce qui est des contraintes sur les déterminants, celles qui portent sur des arguments élémentaires peuvent résulter de la nature du prédicat mais sont le plus souvent indépendantes du mode d'actualisation de celui-ci¹⁸. Par contre, celles qui ont trait aux prédicats peuvent dépendre du verbe support (et donc de son mode d'actualisation) ; ainsi, *faim* n'aura pas la même détermination selon qu'il accepte *avoir* ou *ressentir* comme verbe support :

*Luc a (*de la + *la + une petite + E + très) faim*

*Luc ressent (?de la + la + une petite + *E + *très) faim*

C'est pourquoi nous proposons de décomposer l'actualisation en deux étapes. Nous verrons qu'une telle décomposition est d'un intérêt certain dans une perspective comparatiste. La première étape aura pour tâche de spécifier les particularités des arguments tant du point de vue prépositionnel que de celui de leurs déterminants. Ainsi, pour transformer en phrase grammaticalement acceptable le schéma de phrase ci-dessous :

$P_{(12)} \rightarrow$ boire (homme, bière)

on spécifie tout d'abord les configurations de déterminants que *homme* et *bière* acceptent en tant que, respectivement, sujet et complément du prédicat *boire*. Cela donne la représentation suivante :

$P_{(12)}$ ($D_{\text{ét}1}$ [homme] boire $D_{\text{ét}2}$ [bière])

où les indices $D_{\text{ét}1}$ et $D_{\text{ét}2}$ indiquent les ensembles de déterminants respectivement compatibles avec *homme* et *bière*. De même, pour les formulations suivantes :

$P_{(13)}$ → mécontent (homme, bière)

$P_{(14)}$ → indifférence (homme, bière)

on aboutit respectivement à :

$P_{(13)}$ → $D_{\text{ét}1}$ [homme] mécontent $Pr_{\text{ép}1}$ $D_{\text{ét}3}$ [bière]

$P_{(14)}$ → $D_{\text{ét}1}$ [homme] indifférence $Pr_{\text{ép}2}$ $D_{\text{ét}4}$ [bière]

où les indices $Pr_{\text{ép}1}$, $Pr_{\text{ép}2}$, $D_{\text{ét}3}$ et $D_{\text{ét}4}$ indiquent respectivement des ensembles de prépositions ($Pr_{\text{ép}1}$ = : *de*, $Pr_{\text{ép}2}$ = : *pour + envers*) et de déterminants compatibles avec le nom *bière* tel qu'il apparaît dans $P_{(13)}$ ou $P_{(14)}$.

Pour ce qui est de la seconde étape, il faut distinguer les cas en fonction de la nature du prédicat :

- si c'est un verbe, aucune indication supplémentaire n'étant nécessaire, on peut générer toutes les phrases grammaticales possibles en recourant à des dictionnaires flexionnels :

(12) *Un homme boit de la bière*

(12') *Un homme boit quelques bières*

(12'') *Des hommes boivent quelques bières*

etc.

- si c'est un adjectif, il suffit de rajouter la copule *être* puis de procéder d'une manière identique :

(13) *Un homme est mécontent de la bière*

(13') *Un homme est mécontent d'une bière*

etc.

- si c'est un nom, il faut d'abord indiquer son verbe support et les contraintes sur les déterminants, par exemple :

$P_{(14)}$ → $D_{\text{ét}1}$ [homme] $éprouver$ $D_{\text{ét}5}$ [indifférence] $Pr_{\text{ép}2}$ $D_{\text{ét}4}$ [bière]

on génère ensuite les phrases grammaticalement acceptables :

(14) *Un homme éprouve de l'indifférence pour la bière*

(14') *Un homme éprouve de l'indifférence envers la bière*

etc.

Dans le cas où différents supports standard sont possibles (ainsi que d'éventuelles contraintes sur les déterminants dépendant de la nature de ces supports), on spécifie les différentes configurations possibles. Ainsi, pour le nom *peur* dont la détermination dépend de ses verbes supports (*éprouver* ou *avoir*), on procède de la façon suivante :

$P_{(15)} \rightarrow$ peur (homme)

$P_{(15)} \rightarrow_{\text{Dé}1}$ [homme] peur

$P_{(15)} \rightarrow_{\text{Dé}1}$ [homme] _{éprouver Dé16} peur/_{Dé1} [homme] _{avoir Dé7} peur/

afin d'aboutir aux phrases grammaticales suivantes :

(15) *Un homme éprouve (de la + *E) peur*

(15') *Un homme a (*de la + E) peur*

Le dispositif que nous venons de présenter permet de rendre explicites les contraintes sur les déterminants propres à chacune des deux catégories syntaxiques de substantifs (arguments et prédicats). Un autre de ses avantages, c'est de pouvoir traiter certains des problèmes de détermination liés à l'aspect et à l'expression de l'intensité. Avant d'en rendre compte, nous précisons ce que nous entendons par phrases grammaticales standard.

Parmi toutes les phrases simples qu'il est possible de générer à partir d'une phrase schématisée, nous distinguons celles qui relèvent d'une description standard de toutes les autres. Par description standard, nous entendons description non marquée. Pour les phrases simples non standard, nous ne prenons en compte que trois types de marques :

- des marques stylistiques ;
- des marques aspectuelles ;
- des marques d'intensité.

Pour ce qui est des premières, il s'agit de variantes des verbes support standard¹⁹. Ainsi, *nourrir* est considéré comme une variante stylistique de *avoir* ou *éprouver* dans l'exemple suivant :

(16) *Luc (a + éprouve + nourrit) une certaine tendresse pour Léa²⁰*

En ce qui concerne les deux dernières, nous avons déjà évoqué celles qui ont trait à l'aspect. Les marques d'intensité concernent essentiellement des prédicats relatifs à des états et des événements. Prenons le cas du nom *faim* et de sa forme adjectivale associée, *affamé* :

(17) *Luc a faim*

(18) *Luc est affamé*

Les énoncés ci-dessus sont des énoncés standard. Différents procédés linguistiques permettent de marquer intensivement le prédicat, entre autres :

- l'adjonction de certains déterminants

(17') *Luc a une faim de loup*

- l'insertion d'adverbes :

(18') *Luc est terriblement affamé²¹*

Jusqu'à présent, notre présentation de la phase d'actualisation des prédicats permettait uniquement de générer des phrases grammaticales standards, c'est-à-dire non marquées (dans le sens précisé ci-dessus). Maintenant, nous envisageons la possibilité

de procéder à la génération de phrases aspectuellement marquées. Considérons les énoncés suivants :

- (19) *Luc donne un conseil à Max*
 (19') *Luc donne plusieurs conseils à Max*
 (19'') *Luc multiplie les conseils à Max*

La phrase schématisée suivante :

$P_{(19)}$ conseil \rightarrow Luc, Max)

ne peut générer que (19) dans la mesure où c'est la seule qui est standard, les deux autres étant aspectuellement marqués (aspect fréquentatif). Si l'on admet le point de vue de Z. S. Harris évoqué ci-dessus, c'est-à-dire que l'aspect est un opérateur de phrase, il suffit de rajouter cet opérateur au schéma initial et de le traiter au moment de l'actualisation du prédicat²². Ainsi, si l'on caractérise aspectuellement $P_{(19)}$ (par exemple l'aspect fréquentatif, symbolisé par $ASP^{fré}$), on procédera comme suit :

$ASP^{fré} (P_{(19)} \rightarrow \text{conseil (Luc, Max)})$

Tout d'abord, on procède à l'actualisation des arguments dans la mesure où celle-ci est indépendante de l'aspect du prédicat :

$ASP^{fré} (P_{(19)} \rightarrow [\text{Luc}] \text{conseil}_{Prép3} [\text{Max}])$

Ensuite, on actualise le prédicat en tenant compte de l'opérateur aspectuel spécifié :

$ASP^{fré} (P_{(19)}) \rightarrow [\text{Luc}]_{donner Dét8} \text{conseil}_{Prép3} [\text{Max}]/$
 $[\text{Luc}]_{multiplier Dét9} \text{conseil}_{Prép3} [\text{Max}]/$
 $[\text{Luc}]_{donner ADV1 Dét8} \text{conseil}_{Prép3} [\text{Max}]$

L'indice **Dét8** correspond à un ensemble de déterminants aspectuellement homogènes (*beaucoup de, divers, plusieurs, une série de*, etc.). L'indice **ADV1** indique un ensemble d'adverbes de fréquence. À la suite de ces spécifications, il est possible de générer des phrases comme (19'), (19'') ou :

(19''') *Luc donne souvent un conseil à Max*

Nous constaterons dans la seconde partie l'utilité de ce formalisme pour la traduction automatique.

Une conséquence directe de notre dispositif, c'est qu'il faut multiplier les observations sur les déterminants dans la mesure où il met en avant la dimension lexicale de la détermination. Nous montrons à présent en quoi la notion de classe d'objets est d'une importance capitale pour atteindre cet objectif qui est l'une des principales tâches du projet **DétTAL**.

1.3. Classe d'objets et détermination²³

Après une rapide présentation de la notion de classe d'objets, nous discutons des classes d'objets en tant que postes d'observation des contraintes sur les déterminants.

1.3.1 la notion de classe d'objets²⁴

Les classes d'objets peuvent se concevoir comme une formalisation de la sémantique prenant appui sur les caractéristiques syntaxiques du lexique. Il s'agit d'ensembles d'unités lexicales sémantiquement homogènes que l'on regroupe en classes d'arguments élémentaires, d'une part, en classes de prédicats, d'autre part. Les premières sont constituées uniquement de substantifs; citons à titre d'exemple les <aliments>, les <machines>, les <voies de communication>. Les secondes sont étudiées en fonction de leur catégorie grammaticale²⁵; elles comportent, entre autres les <délits>, <les opérations techniques>, les <sentiments>.

Une classe d'objets est définie principalement par des opérateurs appropriés, c'est-à-dire des verbes, des adjectifs ou des noms qui ont parmi leurs arguments les éléments de la classe qu'ils caractérisent. Parmi les opérateurs appropriés, on distingue ceux qui s'appliquent à tous les items de ceux qui ne concernent qu'un sous-ensemble de ces items; les premiers sont appelés des opérateurs définitionnels. Ainsi, les adjectifs *brillant*, *mat*, *terne*, *vif* sont des opérateurs appropriés définitionnels de la classe des noms de <couleurs> puisqu'ils s'appliquent à tous les substantifs de la classe alors que les adjectifs *sombre* ou *clair* sont des opérateurs appropriés non définitionnels dans la mesure où ils ont une portée distributionnelle moindre:

*ce blanc est (brillant + *clair + *sombre + mat + terne + vif)*

ce bleu est (brillant + clair + sombre + mat + terne + vif)

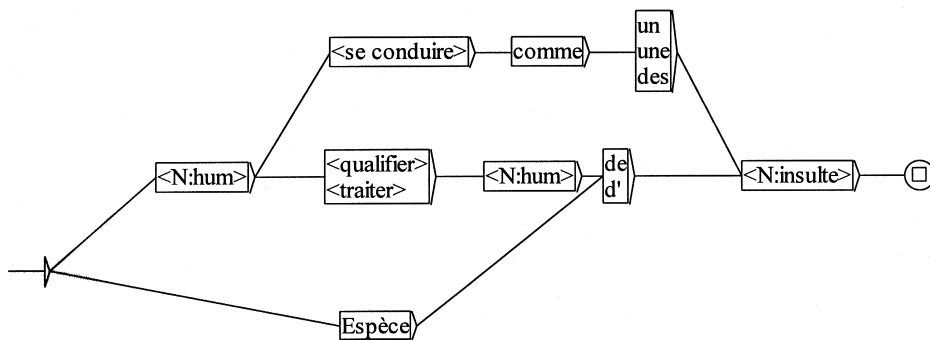
*ce noir est (brillant + *clair + *sombre + mat + terne + vif)*

*ce rouge est (brillant + clair + *sombre + mat + terne + vif)*

ce vert est (brillant + clair + sombre + mat + terne + vif)

Les classes de prédicats nominaux peuvent être également définies par des verbes supports appropriés. Ainsi, *commettre* et *perpétrer* sont des supports caractéristiques de la classe des <crimes et délits>.

Élaborer une classe d'objets revient à construire une grammaire locale autour des éléments constitutifs de cette classe dans la mesure où il ne s'agit pas uniquement de recenser les unités lexicales que subsume le nom de la classe mais également d'indiquer leur syntaxe. Une grammaire locale peut être représentée formellement par un automate à états finis (sous forme de graphe). Ainsi, il est possible de rendre compte de la classe <insultes> de la façon suivante²⁶:



Graphe 1

1.3.2 *Études des contraintes sur les déterminants au regard des classes d'objets*

Nous proposons de décrire les contraintes sur les déterminants en ayant recours aux classes d'objets dans la mesure où ces dernières, en tant que descripteurs exhaustifs du lexique, permettent de systématiser les observations sur les conditions d'apparition d'un déterminant en fonction des particularités syntaxiques du nom qui l'accepte. Le recours méthodique aux classes d'objets permet de rapprocher des faits de détermination alors qu'une approche au cas par cas le permettrait plus difficilement et qu'une analyse ne tenant pas compte de la dimension lexicale de la détermination l'interdit *de facto*. Les faits de détermination ainsi regroupés peuvent concerner aussi bien les modificateurs libres des noms que les autres éléments déterminatifs; toutefois, les opérations de regroupement ne sont pas de même nature selon que l'on traite le premier type de phénomènes ou le second.

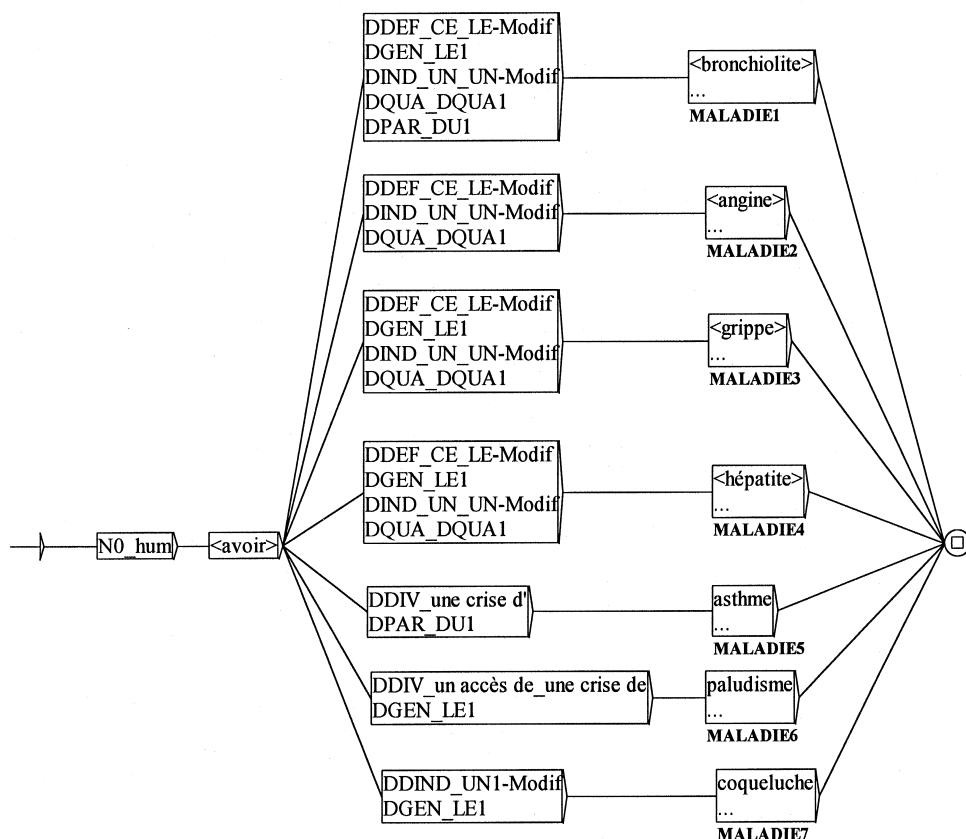
Dans les cas où la modification libre d'un nom est assimilable à une forme réduite de prédication, les classes permettent de rendre compte de la compatibilité entre les modificateurs et les noms dans la mesure où tout prédicat approprié exprimable en termes de relative est un candidat modifieur sous une forme propre à ce type de construction²⁷. Pour reprendre l'exemple des noms de <couleurs> évoqué ci-dessus, on déduira de la description de cette classe que tous les noms de couleurs admettent une classe de modificateurs adjectivaux :

un (blanc + bleu + rouge + noir + ...) (brillant + mat + terne + vif)

et qu'une partie d'entre eux seulement sont compatibles avec une autre classe de modificateurs adjectivaux :

*un (*blanc + bleu + rouge + *noir + ...) (clair + sombre)*

En dehors de la modification, il s'agit de faire apparaître des sous-ensembles de noms relatifs à telle ou telle combinatoire avec les déterminants en construisant les grammaires locales relatives à chaque classe d'objets. Les représentations sous forme de graphes permettent de formaliser les différents types de contraintes sur les déterminants qui caractérisent les substantifs d'une classe donnée. Ainsi, pour ce qui est des noms de maladies précédés du support *avoir*, on obtient le graphe suivant (il s'agit d'un extrait) :



Graphe 2

Dans le cadre du projet DétTAL, nous multiplions les observations relatives aux contraintes sur les déterminants en utilisant le formalisme des classes d'objets.

Nous considérons désormais que la description de la détermination de chacune des deux langues s'effectue sur les bases formelles que nous venons d'évoquer. Nous nous intéressons à présent au transfert des contraintes sur les déterminants d'une langue source vers une langue cible²⁸.

2. LES MODULES DE TRADUCTION

Notre réflexion sur la traduction des déterminants n'est évidemment pas indépendante de celle sur la traduction des autres unités linguistiques. Tout projet en TA (Traduction Automatique) se doit d'avoir recours à des dictionnaires électroniques à large couverture, et ce, pour chacune des deux langues mises en regard. Nous ferons donc d'abord quelques observations concernant l'élaboration de la description linguistique dans les dictionnaires. Ensuite, nous considérerons des applications de ces descriptions à la traduction automatique espagnol-français.

2.1. Les dictionnaires monolingues coordonnés

Tout le monde s'accorde pour dire que le dictionnaire doit fournir un maximum d'informations sur les unités qu'il présente (les lemmes). Le dictionnaire bilingue se doit en plus de nous renseigner sur le comportement linguistique des équivalents de traduction qu'il propose, comportement qui ne peut pas être assimilé à celui des lemmes par le seul fait qu'il existe entre eux une relation d'équivalence de traduction. En effet, la description suivante d'un opérateur approprié à la classe des <délits> :

castigar/N0:hum/N1:hum/N2:con <peine>/N3:
por <délit>/Fr:punir²⁹

ne peut fonctionner à elle seule comme module de traduction, puisque des informations particulières sur *punir* ne sont pas directement déduisibles du lemme espagnol (ainsi, par exemple, *punir* régit la préposition *de* pour son deuxième complément et non pas *avec*). Un dictionnaire électronique doit être totalement explicite et ne peut pas avoir recours à des exemples que le dictionnaire bilingue sur papier emploie pour rendre compte de ce type de différences entre L1 et L2. Il faut donc concevoir le module de traduction non comme un dictionnaire bilingue, mais comme un système de pointeurs d'un dictionnaire monolingue à un autre³⁰. Le contenu du champ *Fr* ci-dessus nous aiguillerait vers :

punir/N0:hum/N1:hum/N2:de <peine>/N3:
pour <délit>/Es:castigar³¹

Pour un nom prédicatif comme *contrabando*, on aurait :

contrabando/M:n1/N0:hum/N1:de
Ninc/W:hacer/Dét:E/Fr:contrebande
contrebande/M:n21/N0:hum/N1:de
Ninc/W:faire/Dét:DU/Es:contrabando³²

Nous appelons « coordonnées » ces descriptions couplées par une relation de traduction. Elles permettent de rendre compte de façon économique des éventuelles différences quant aux propriétés linguistiques des unités dans chaque langue.

Les cas de polysémie du lemme conduisant à différents équivalents de traduction seraient résolus grâce au dédoublement des entrées en L1 :

secuestro/N0:hum/N1:hum/Fr:enlèvement
secuestro/N0:hum/N1:<moyen de transport en commun>/Fr:détournement
violación/N0:hum/N1:hum/Fr:viol
violación/N0:hum/N1:<loi>/Fr:violation

Le choix entre différents synonymes en langue d'arrivée dépendrait du système de traitement synonymique du dictionnaire de L2. Ainsi, c'est au dictionnaire français d'opérer éventuellement le choix entre *enlèvement*, *kidnapping*, *kidnappage*...

Le module de traduction assure donc le passage d'une entrée en L1 à une entrée en L2. Ce passage ne s'effectue que pour les prédicats et leurs arguments, jamais pour les moyens de linéarisation ou d'actualisation de ceux-ci. Ainsi, par exemple, les prépositions en tant que marques d'arguments ne peuvent pas, à proprement parler, être traduites. Elles doivent être fournies par la description lexicographique associée au prédicat de L2. Le même principe s'applique aux verbes supports.

2.2. Vers une traduction automatique des déterminants

Nous avons vu que la détermination présente des différences notables selon qu'elle concerne des arguments élémentaires ou des noms prédicatifs.

La traduction d'un déterminant associé à un nom prédicatif est strictement dépendante de l'analyse que l'on propose pour le déterminant. En schématisant, un déterminant peut fonctionner comme :

- un actualisateur du nom prédicatif, qui peut être aspectuellement neutre ou marqué ;
- un quantifieur³³, qui peut introduire une quantification discrète ou approximative.

2.2.1 Le déterminant comme actualisateur aspectuellement neutre

Si le déterminant est un actualisateur neutre du nom prédicatif, il est codé dans nos dictionnaires d'après la classe syntactico-sémantique du nom prédicatif en question. Ainsi, on aura :

Juan (tiene + padece de) E (anorexia + asma)
Juan ha tenido EL (peste + cólera)
Juan (tuvo + sufrió) UN (infarto + cólico hepático)

En dépendant de la reproductibilité des intuitions sémantiques et de la fiabilité des tests syntaxiques développés pour le classement des noms, on pourra soit mettre en rapport le comportement vis-à-vis de la détermination avec l'appartenance du nom à une sous-classe sémantico-aspectuelle donnée (par exemple les noms de maladies épidémiques), soit se contenter de coder en extension, à l'intérieur d'une classe donnée, le déterminant par défaut correspondant à chaque nom. Rappelons que le déterminant peut dépendre de facteurs comme le choix du verbe support :

*Juan practica el fútbol (Dét = : EL) (*Juan practica fútbol)*
Juan juega (al + a) fútbol (Dét = : EL+E)
*Juan hace fútbol (Dét = : E) (*Juan hace el fútbol)*

Or, a priori, le sens de *fútbol* dans ces trois phrases est le même. Les variations dans la détermination ne peuvent être mises en rapport ici avec la sémantique³⁴.

En tout état de cause, ce type de détermination neutre, purement au service de l'actualisation, n'implique pas la transmission d'un contenu sémantique propre au déterminant lors de la traduction. Le nom prédicatif sera donc associé à son équivalent de traduction qui, lors de la mise en discours, recevra son déterminant « par défaut » du dictionnaire monolingue de la langue d'arrivée.

Considérons la phrase :

Juan tiene asma

Le prédicat nominal comporte un verbe support standard (sans marques stylistiques, aspectuelles ou intensives) et le déterminant (dans ce cas, zéro) ne comporte pas de charge sémantique (quantitative ou métadiscursive) particulière. Les seules relations de traduction à considérer seront celles qui correspondent au schéma d'arguments *asma(Juan)*. Le module en langue d'arrivée sera chargé de générer la traduction à partir du schéma *asthme(Jean)* :

Jean a de l'asthme

Le partitif est employé parce qu'il est l'actualisateur standard du nom prédicatif asthme, sans qu'il soit nécessaire de postuler *in abstracto* que le déterminant zéro espagnol correspond au partitif français, ce qui constituerait, à cette étape, une généralisation fort risquée et, somme toute, superflue.

De la même façon, des phrases comme :

El ministro cometió prevaricación

El ministro incurrió en prevaricación

seront traduites à partir du nom prédicatif :

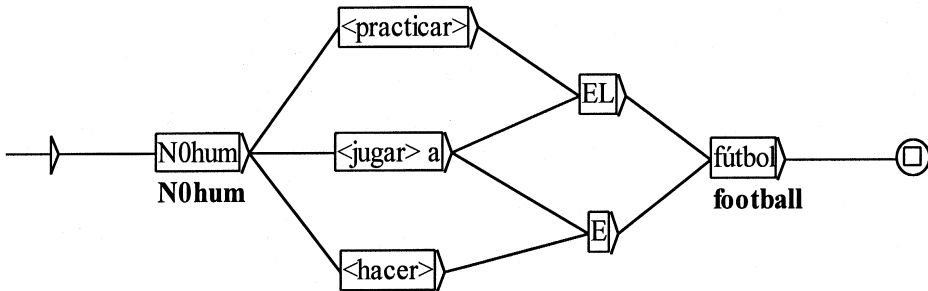
prevaricación/N0:hum/Fr:forfaiture

Les verbes *cometer* et *incurrir en* sont des supports possibles pour *prevaricación*. Ils comportent le déterminant zéro. Le français *forfaiture* sélectionnera le support *commettre*, qui comportera par défaut l'indéfini UN³⁵ :

Le ministre a commis une forfaiture

Il est important de souligner que, même si la solution que la langue d'arrivée adopte pour un cas concret est formellement identique à celle employée par la langue de départ (ce qui arrive souvent entre des langues voisines), le procédé de passage de L1 à L2 est modélisé de la même façon, ce qui permet de maintenir un haut degré d'indépendance des descriptions monolingues par rapport au module bilingue concernant une paire de langues données.

Précisons, par ailleurs, qu'aussi bien les informations nécessaires à la reconnaissance en langue de départ qu'à la génération en langue d'arrivée sont représentées à partir d'automates à états finis qui rendent compte des différentes contraintes sur les déterminants pour une classe donnée de noms prédicatifs. Cette représentation semble nécessaire dans la mesure où des contraintes croisées existent entre la détermination d'un nom prédicatif et ses verbes supports. Une représentation linéaire ne serait donc pas opérationnelle, car elle obligerait à dédoubler des entrées non à partir de la sémantique, mais de l'actualisation. Pour l'exemple de *fútbol* nous aurons³⁶ :



Graphe 3

2.2.2 Les déterminants et l'expression de l'aspect

Si le verbe support véhicule un contenu aspectuel, le principe de passage de L1 à L2 est similaire (calcul en langue d'arrivée), étant donné que l'expression de l'aspect avec les noms prédicatifs n'est pas régulière, mais fortement collocationnelle. Chaque nom prédicatif sélectionne des verbes supports aspectuels appropriés, qui comportent un déterminant spécifique, souvent différent de celui qui accompagne le verbe support standard. Il est toutefois nécessaire d'indiquer, dans le schéma d'arguments qui sera traduit, l'aspect de la phrase en L1. Ainsi, par exemple :

Juan se da a la fuga

sera traduit à partir de **fuga(Juan)** plus indication que *fuga* a un aspect inchoatif, c'est-à-dire $ASP^{incho}(fuga(Juan))$. La traduction, à proprement parler, serait $ASP^{incho}(fuite(Jean))$. La valeur inchoative pour *fuite*, calculée en langue d'arrivée, correspondrait à *prendre la* (et non pas *se donner à la*) :

Jean prend la fuite

Le déterminant peut introduire à lui seul une valeur aspectuelle (par exemple itérative). Le procédé à adopter serait le même.

Juan soltó una retahila de tacos

implique $ASP^{itér}(taco(Juan)) = ASP^{itér}(injure(Jean))$:

Jean proféra (un chapelet + une bordée) d'injures

Étant donné que l'aspect peut être véhiculé soit par un verbe support, soit par un déterminant, soit par un adverbe, rien n'empêche de considérer que la langue d'arrivée exprime l'aspect ailleurs que dans la langue de départ. Le procédé présenté permet de rendre compte de façon naturelle de ces cas :

Últimamente, Juan ha hecho un montón de tonterías

Ces derniers temps, Jean a multiplié les bêtises

Tu lui as reflaqué une claque

Le has pegado otra bofetada³⁷

Il suffit que la représentation monolingue moyennant des automates indique quels sont les moyens les plus naturels pour exprimer le fréquentatif ou l'itératif.

Notons qu'il faudrait probablement accorder une place un peu particulière à l'expression de l'intensité. De la même façon, l'expression de l'aspect a recours souvent à des modificateurs, qui seraient à étudier dans le cadre de la détermination (*UN-Modif*)³⁸.

2.2.3 Les déterminants comme quantifieurs discrets

Dans certains cas, le déterminant fonctionne comme un prédicat quantifieur. Les cardinaux et les déterminants nominaux indiquent, normalement³⁹, une quantité discrète. Ils peuvent être traduits de façon directe. Notons que la quantification s'applique aussi bien à des arguments élémentaires qu'à des prédicats :

Juan hizo tres preguntas
tres:pregunta, pregunta(Juan)
trois:question, question(Jean)
Jean posa trois questions

Juan compró dos caramelos
dos:caramelo, comprar(Juan, caramelo)
deux:bonbon, acheter(Jean, bonbon)
Jean a acheté deux bonbons

Pour traiter ces cas, le concept de classe d'objets s'avère un outil efficace aussi bien pour les arguments que pour les prédicats. Notamment, il nous semble qu'une typologie sémantique des noms prédicatifs soit essentielle pour délimiter les frontières imprécises entre aspect et quantification dans le cas de la détermination de ces noms.

2.2.4 Le déterminant comme quantifieur approximatif

Dans d'autres cas, la quantification est approximative. Un travail comparatif spécifique semble nécessaire pour traiter ces cas, pour lesquels on peut adopter une solution mixte qui tienne compte de la forme du déterminant de L1, mais qui prévoie des passages par des classes sémantiques. *A priori*, le système pourrait marcher entre l'espagnol et le français à partir des formes.

Pour un prédicat :

Juan ha hecho demasiadas lecturas
demasiado:lectura, lectura(Juan)
trop:lecture, lecture(Jean)
Jean a fait trop de lectures

Pour un argument :

Juan ha leído demasiados libros
demasiado:libro, leer(Juan, libro)
trop:livre, lire(Jean, livre)
Jean a lu trop de livres

Mais la robustesse du système peut s'en ressentir, les formes en principe équivalentes dans les deux langues ne présentant pas forcément un fonctionnement identique. Il serait donc souhaitable de disposer d'une typologie de quantifieurs qui permette de faire passer au système une indication de type sémantique. Le déterminant serait alors généré en langue d'arrivée. Par exemple, pour la quantification dite « moyenne⁴⁰ » :

Juan ha leído bastantes libros
QUANT^{moy}:libro, leer(Juan, libro)
QUANT^{moy}:livre, lire(Jean, livre)
Jean a lu (pas mal + assez) de livres

Des indications diastatiques pourraient être couplées à la typologie de quantification. Ainsi, nous trouvons de nombreuses formes familières pour la « quantification forte absolue » :

Juan ha enviado un huevo de archivos (por correo electrónico)
QUANT^{fortabs_fam}:archivo, enviar(Juan, archivo)
QUANT^{fortabs_fam}:fichier, envoyer(Jean, fichier)
Jean a envoyé une chiée de fichiers (par e-mail)

Un traitement semblable pourrait être envisagé pour des cas plus épineux, comme l'emploi des articles comme quantifieurs universels ou des formes comme *des* en tant que quantifieur approximatif.

La typologie comparée des quantifieurs espagnol-français est actuellement en cours. On observe des trous, comme par exemple la quantification « très forte axiologique » en français (*beaucoup trop de livres*), qui n'a pas d'expression régulière en espagnol ou la quantification « très faible axiologique négative » (*trop peu d'eau*) qui pose des problèmes en espagnol (?*demasiada poca agua; demasiado poca*). Cependant, les deux langues restent assez proches entre elles⁴¹.

2.3. Le choix en langue cible

Le procédé de traduction indiqué laisse un certain espace au choix de formes en L2. Cela implique qu'il faut développer un système de choix de la solution finale adoptée (même si, *a priori*, on postule une certaine équivalence des solutions possibles), mais présente en même temps un avantage considérable à nos yeux : il n'existe pas des zones de L2 « inaccessibles » à partir de L1, ce qui confère à la traduction une souplesse toute particulière, même en TA.

Rappelons, en effet, que de nombreux travaux en traductologie abordent la question des divergences de traduction. Or, ils ont tendance à ne considérer que les divergences obligatoires, alors que le traducteur (nous serions tenté de dire : le « bon » traducteur) crée constamment, par rapport au texte de départ, des divergences « optionnelles ». C'est-à-dire qu'il procède au choix de solutions plus conformes à l'esprit de la langue d'arrivée en dépit de solutions plus proches de la langue de départ, grammaticalement acceptables en langue d'arrivée, mais senties comme « moins bonnes ». Qui plus est, le traducteur fait varier les solutions adoptées en fonction de sa perception du texte en langue d'arrivée.

Reprenons l'exemple :

Jean joue au football

**Jean joue à football*

Juan juega al fútbol

Juan juega a fútbol

La phrase *Juan juega a fútbol*, acceptable en espagnol, serait « inaccessible » à partir du français si l'on traduit en s'attachant aux formes, car le passage s'effectuerait toujours de *Jean joue au football* à *Juan juega al fútbol*, (**Jean joue à football*) étant impossible en français.

Si le système traduisait à partir du catalan, la situation serait la suivante :

**En Joan juga al futbol*

En Joan juga a futbol

Juan juega al fútbol

Juan juega a fútbol

Dans ce cas, ce serait la solution *Juan juega a fútbol* qui serait systématiquement retenue au détriment de *Juan juega al fútbol*, qui resterait « inaccessible ».

Par contre, si la traduction est générée en L2 à partir de *fútbol*(*Juan*), les deux formes sont également possibles et son choix peut être dirigé par des critères internes

à L2 que l'on peut implémenter indépendamment de la langue de départ. Le système est, dans ce sens, plus robuste car plus à même de simuler le maniement de la langue par un traducteur humain et moins dépendant de la langue de départ.

Soulignons qu'il ne s'agit pas là de cas rares ou de curiosités. Cet état de choses touche à des pans importants du système de la langue. Pensons, par exemple, aux intensifs élatifs (formes en *-ísimo*), courants en espagnol, mais rares en français.

3. EN GUISE DE CONCLUSION

La traduction doit être pilotée non par des formes, mais par les sens ou par le rôle syntactico-sémantique de certaines occurrences des formes de la langue de départ. Ce principe, qui nous semble valable de façon générale en traduction, et *a fortiori* en TA, s'avère particulièrement pertinent pour la traduction des déterminants où les rapports forme/sens-fonction sont fortement asymétriques.

La traduction implique une analyse suffisamment fine de la langue de départ pour être en mesure de choisir la bonne stratégie de traitement. Cette analyse implique, pour le sujet qui nous occupe, le développement de descriptions linguistiques précises des différents moyens d'actualisation des noms prédicatifs, ce qui soulève au moins deux questions importantes :

- pour les noms prédicatifs : il faut disposer, pour chaque langue, d'une typologie syntactico-sémantique des noms prédicatifs couplée aux différents verbes supports acceptés par ceux-ci⁴² et aux moyens d'expression des différentes valeurs aspectuelles ;
- pour les noms prédicatifs et les arguments élémentaires : il faut disposer d'une typologie de la quantification approximative suffisamment précise pour garantir le passage d'une forme en L1 à une forme en L2 à travers une étiquette de type sémantique. Notons qu'une structuration syntactico-sémantique du lexique est également nécessaire pour les arguments élémentaires⁴³.

NOTES

1. D'autres langues, dont le grec, l'allemand et l'anglais, sont prévues par la suite.
2. Dans le domaine plus général de la linguistique, il y a de nombreux travaux sur la détermination. Pour ce qui est du TALN, les avancées les plus significatives concernent essentiellement l'analyse automatique où le problème de la détermination est relativement simple à résoudre sur des bases morpho-syntaxiques.
3. Les travaux sur la quantification et les descriptions définies de la philosophie analytique sont à l'origine de la plupart de ces modèles. Cf. Corblin (1987), Corblin (1995), Kleiber (1981), Kleiber (1990) ou Milner (1978) pour le français.
4. C'est-à-dire *le(la)*, *les*, *un(e)* en tant que formes de la quantification universelle. Contrairement à Corblin (1987), nous ne pensons pas que la liste des génériques comporte d'autres formes.
5. Cf. Gross, M. (1986).
6. Nous avons contribué à ce recensement, cf. Buvet (1993) pour le français et Blanco (1997) pour l'espagnol.
7. Plus de 3000 si l'on tient compte de tous les déterminants nominaux, sinon une centaine.
8. Sur la typologie des déterminants, cf. Buvet (à paraître a).
9. Cf. Gross, M. (1981), Gross, G. et Vivès (1986). Signalons que le nombre et la nature des arguments dépendent de chacun des prédicats.
10. Giry-Schneider (1979), Giry-Schneider (1987), Gross, G. (1989), Meunier (1981) et Vivès (1983).
11. Sur la détermination des noms de maladie, cf. Buvet (1998) pour le français et Blanco et Bonell (à paraître) pour l'espagnol.

12. Cf. Lim (1998). L'inchoatif est également un autre aspect que peuvent indiquer les déterminants : *Luc avait un début de grippe*.
13. Cf. Gross, G. (1989).
14. Cf. Lim (1998)
15. P, Préd, Arg, symbolisent respectivement une phrase, un prédicat et un argument (au minimum). Les positions syntaxiques des arguments sont spécifiées par des numéros : Arg₀, Arg₁, Arg₂, Arg₃, qui correspondent respectivement au sujet et aux premier, second et troisième compléments du prédicat.
16. Remarquons que dans les cas où le prédicat est un substantif le verbe support n'apparaît pas dans la représentation schématisée de la phrase. Le statut d'actualisateur de ce type de verbe explique cette absence.
17. De nombreux travaux relevant de cette représentation de la phrase simple sont listés dans Leclere (1998).
18. Le nombre de situations où le mode d'actualisation du prédicat influe sur la détermination des arguments serait très limité. Signalons le cas suivant : (*De l' + l') *argent m'intéresse*, (De l' + *l') *argent m'intéresserait*.
19. Cf. Gross, M. (1998).
20. La possibilité pour un support d'avoir telle ou telle variante est subordonnée à la nature du prédicat nominal ; cf. Gross, G. et Vivès (1986).
21. Des verbes supports sont également possibles pour certains prédicats nominaux, cf. *Luc explose de joie*. Signalons également que l'intensité peut être forte (*Luc a beaucoup d'estime pour Tom*) ou faible (*Luc a peu d'estime pour Tom*) ; cf. Buvet et Lim (1996).
22. Sur les rapports entre l'actualisation et l'aspect, cf. Gross, G. (1996).
23. Sur les rapports entre classes d'objets et détermination, cf. Buvet (1998), Buvet (à paraître b).
24. Cf. Gross, G. (1992), Gross, G. (1995), Le Pesant et Mathieu-Colas (1998) et Mathieu-Colas (1998).
25. Si un prédicat a différentes formes, les liens existant sont spécifiés dans les dictionnaires *ad hoc*.
26. La grammaire locale des <insultes> est d'une grande simplicité par rapport aux autres grammaires locales correspondant à des classes de prédicats ; cf. Gross, M. (1995) à propos de la grammaire locale des noms de sentiments qui est nettement plus complexe.
27. Les compatibilités avec les prédicats non appropriés résultent d'héritages divers que l'on précise en définissant la classe (cf. Le Pesant (1998)).
28. Du transfert de l'espagnol au français, d'une part, et de celui du français vers l'espagnol, d'autre part. Chacun des deux transferts est conçu d'une manière indépendante.
29. Nous nous servons de la notation couramment employée dans les travaux du LLI (les champs *Ni* renvoient à un équivalent de traduction en français du lemme espagnol). Sur le codage des dictionnaires électroniques du LLI, cf. Mathieu-Colas (1994).
30. Cela permet aussi de régler des cas triviaux mais qui engendrent, dans les dictionnaires bilingues traditionnels, un certain degré de redondance, par exemple la divergence de genre entre un nom en L1 et son équivalent de traduction en L2, par exemple *fraude* n.m. (esp.) = *fraude* n.f. (fr.) ou des différences quant à la flexion en nombre : *atraco* (m.p. : *atracos*) = *hold-up* (invariable).
31. Le champ *Es* renvoie à un équivalent de traduction en espagnol du lemme français.
32. Cette représentation est une simplification, car les contraintes croisées existant entre verbe(s) support(s) (champ *W*) et déterminant(s) obligent souvent à avoir recours à des automates à états finis pour la représentation de l'actualisation du nom prédicatif. À ce moment, le passage ne s'effectue pas entre deux entrées de dictionnaire sous forme linéaire, mais entre des graphes (cf. *infra*).
33. Les déterminants peuvent être aussi de nature métadiscursive (par exemple marqueurs de corréférence). Dans ce cas (déterminants définis), la traduction exige la prise en compte de facteurs qui vont au-delà de la phrase simple. Nous n'aborderons pas cette question dans le cadre de cet article.
34. Précisons que, malgré la présence de variantes, il existe une relation appréciable entre les classes sémantiques des noms prédicatifs et leurs verbes supports (ainsi, les <délits> se combinent avec *commettre*, les <sports> avec *pratiquer*, les <maladies> avec *avoir*, etc.). De là l'utilité des classes d'objets comme cadre pour la codification d'informations concernant la détermination.
35. Cf. les contextes suivants extraits, pour l'espagnol, du journal *El Mundo* et, pour le français, du *Monde* : « Villar y Valenzuela incurrieron en prevaricación, al conceder 135 millones a Renasa » ; « ... señala que Pérez Villar cometió prevaricación al dictar una resolución injusta a sabiendas » ; « ... sous peine de « commettre une forfaiture ». Une nouvelle procédure? ».

36. Le nom prédicatif *fútbol* partagerait cette partie de la description linguistique avec des noms comme *baloncesto*, *tenis*, ce qui permet un certain degré de compactage de l'information. Par contre, des noms comme *natación*, *atletismo* n'admettraient que les supports *practicar* et *hacer*.
37. Remarquons l'impossibilité suivante : **le has remetido una bofetada*.
38. Cf. Blanco, Buvet et Gavriilidou (à paraître).
39. Sauf dans des emplois métaphoriques, figuratifs ou hyperboliques. Dans ces cas, ils sont soit des aspectuels, soit des intensifs, soit encore des collectifs (Blanco (1998)). Ces trois cas de figure impliquent l'existence d'une collocation et, donc, l'impossibilité d'une traduction proprement dite.
40. Nous nous appuyons sur une typologie de la quantification proposée par Gauchola et Mestreit (inédit). Dans cette section, nous reprenons entre guillemets les termes employés par ces auteurs.
41. Faisons remarquer aussi la frontière imprécise entre la quantification d'un argument donné et l'intensité appliquée au prédicat (relations paraphrastiques du genre : *Jean lit beaucoup de livres* = *Jean lit beaucoup*).
42. Notons aussi qu'il semble nécessaire de prendre en compte le niveau diastatique de l'expression (particulièrement les formes familières), qui semble être étroitement imbriqué à l'actualisation des noms prédicatifs à travers le choix des verbes supports (par exemple *donner une gifle* vs *foutre une gifle*) et de certains déterminants nominaux (par exemple *une chiée*, *une tapée d'arrestations*).
43. Ne serait-ce que pour reprendre la célèbre distinction entre noms comptables et noms massifs. La structuration en classes d'objets nous semble cependant présenter de multiples avantages sur ce premier mode de classement.

RÉFÉRENCES

- BLANCO, Xavier (1997) : « Un dictionnaire électronique des déterminants nominaux en espagnol », *BULAG, numéro spécial Actes du Colloque FRACTAL 97*, Besançon, Université de Franche-Comté, p. 45-54.
- (1998) : « Les déterminants nominaux figés. Perspective contrastive espagnol-français », *Actes des Rencontres linguistiques méditerranéennes*, Tunis, p. 19-34.
- BLANCO, Xavier et Carmen BONELL (1998) : « Vers une structuration syntactico-sémantique de la terminologie médicale. Applications à la traduction français-espagnol », *Cahiers de grammaire*, 23, p. 1-18.
- BLANCO, Xavier, Pierre-André BUVET et Zoé GAVRIILIDOU (à paraître) : « Études comparatives des modificateurs figés ».
- BUVET, Pierre-André (1993) : *Les déterminants nominaux quantifieurs*, thèse de doctorat, Université Paris XIII.
- (1998) : « Détermination et classes d'objets », *Langages*, 131, p. 91-102.
- (à paraître a) : *Remarques sur la détermination du français*.
- (à paraître b) : « La détermination des noms prédicatifs ».
- BUVET, Pierre-André et Jung-Hae LIM (1996) : « Les déterminants nominaux aspectuels », *Linguisticae Investigationes*, p. 271-286.
- CORBLIN, Francis (1987) : *Indéfini, défini et démonstratif : constructions linguistiques de la référence*, Genève/Paris, Droz.
- (1995) : *Les formes de reprise dans le discours – Anaphores et chaînes de référence*, Rennes, PUR.
- GAUCHOLA, Roser et Claude MESTREIT (inédit) : « La quantification en français, espagnol et catalan ».
- GIRY-SCHNEIDER, Jacqueline (1979) : *Les nominalisations en français*, Genève/Paris, Droz.
- (1987) : *Les prédicats nominaux en français*, Genève/Paris, Droz.
- GROSS, Gaston (1989) : *Les constructions converses du français*, Genève/Paris, Droz.
- (1992) : « Forme d'un dictionnaire électronique », *L'environnement traductionnel. La station de travail de l'an 2001*, Sillery, Presses de l'Université du Québec, p. 255-272.
- (1995) : « Une sémantique nouvelle pour la traduction automatique : les classes d'objets », *La tribune des industries de la langue et de l'information électronique*, 17-18-19.
- (1996) : « Prédicats nominaux et compatibilité aspectuelle », *Langages*, 121, p. 54-72.
- GROSS, Gaston et Robert VIVÈS (1986) : « Les constructions nominales et l'élaboration d'un lexique-grammaire », *Langue française*, 69, p. 5-27.

- GROSS, Maurice (1981): « Les bases empiriques de la notion de prédicat sémantique », *Langages*, 63, p. 7-52.
- (1986): *Grammaire transformationnelle du français*, tome 2, Paris, Cantilène.
- (1995): « Une grammaire locale de l'expression des sentiments », *Langue française*, 105, p. 70-87.
- (1998): « La fonction sémantique des verbes supports », *Travaux de Linguistique*, 37, p. 25-46.
- HARRIS, Zellig S. (1976): *Notes du cours de syntaxe*, Paris, Seuil.
- KLEIBER, Georges (1981): *Problèmes de référence: descriptions définies et noms propres*, Paris, Klincksieck.
- (1990): *L'article LE générique. La généricité sur le mode massif*, Genève/Paris, Droz.
- LE PESANT, Denis (1998): « Utilisation de propriétés des anaphores dans la définition des relations lexicales », *Langages*, 131, p. 115-124.
- Le Pesant, Denis et Michel MATHIEU-COLAS (1998): « Introduction aux classes d'objets », *Langages*, 131, p. 6-33.
- LECLÈRE, Christian (1998): « Travaux récents en lexique-grammaire », *Travaux de Linguistique*, 37, p. 155-186.
- LIM, J.-H. (1998): *La fréquence et son expression en français*, thèse de doctorat, Université Paris XIII-LLI, Villetaneuse.
- MATHIEU-COLAS, Michel (1994): *Les mots français à trait d'union*, Paris, Didier Érudition.
- (1998): « Illustration d'une classe d'objets: les voies de communication », *Langages*, 131, p. 77-90.
- MEUNIER, Annie (1981): *Nominalisations d'adjectifs par verbes supports*, thèse de troisième cycle, Université Paris VIII-LADL.
- MILNER, Jean-Claude (1978): *De la syntaxe à l'interprétation*, Paris, Seuil.
- VIVÈS, Robert (1983): *Avoir, prendre, perdre: constructions à verbe support et extensions aspectuelles*, thèse de troisième cycle, Université Paris VIII-LADL.