



HAL
open science

Théories et méthodes dans la comparaison des langues : chemins vers la linguistique générale

Anna Sörös

► **To cite this version:**

Anna Sörös. Théories et méthodes dans la comparaison des langues : chemins vers la linguistique générale. Sciences de l'Homme et Société. Université Paris X Nanterre, 2007. tel-01081421

HAL Id: tel-01081421

<https://shs.hal.science/tel-01081421>

Submitted on 7 Nov 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Anna Sörös

Théories et méthodes dans la comparaison des
langues : chemins vers la linguistique générale

Document de synthèse pour l'HDR

Université Paris X Nanterre
2007

Date de soutenance : 08/11/2007

Directeur : Bernard Laks

Membres du jury : Denis Creissels

Christiane Marchello-Nizia

Bertrand Boiron

Table des matières

1. La diversité des langues dans la linguistique générale
 - 1.1. La linguistique générale
 - 1.2. Les linguistes sur la diversité des langues

2. Sur la comparaison des langues
 - 2.1. Les différents types de comparaisons
 - 2.2. La terminologie

3. Avant les comparaisons, la description
 - 3.1. Pourquoi le hongrois ?
 - 3.3. L'ordre des mots
 - 3.4. La structure informationnelle
 - 3.5. L'agglutination
 - 3.6. Le passif
 - 3.7. La grammaticalisation de l'ESPACE

4. Comparaisons intragénétiqes
 - 4.1. Les langues romanes
 - 4.1.1. Ordre des mots
 - 4.1.2. Et où placer les créoles ?
 - 4.2. Langues finno-ougriennes

5. Comparaisons intergénétiqes. Approches typologiques
 - 5.1. La place de l'adjectif épithète dans les langues
 - 5.2. Les adpositions

6. Comparaisons intergénétiqes. Approches contrastives
 - 6.1. Méthodologie : les articles
 - 6.2. Ordre des mots
 - 6.3. Temps et aspect

7. Comparaisons en diachronie

8. L'évolution de la morphologie des langues romanes : flexionnelle et analytique ?
 - 8.1. Introduction
 - 8.2. Analyse de la technique morphologique du français
 - 8.3. De synthétique en analytique

- 8.4. Les articles
- 8.5. Disparition de la flexion nominale et emploi de prépositions
- 8.6. La gradation des adjectifs
- 8.7. Les pronoms personnels
- 8.8. Les auxiliaires
- 8.9. Synthèse et analyse dans le système verbal
- 8.10. D'analytique en synthétique: Les articles contractés
- 8.11. Conclusions

Conclusions et perspectives

Mémoire de synthèse de Anna Sörös pour l'HDR

Introduction : Hommage à Joseph Herman (1924-2005)

A un moment donné de sa carrière de chercheur, non seulement on fait un bilan de ses activités, mais on pense aussi à ceux qui l'ont guidé pendant cette période. Plus la carrière est longue, plus la liste des personnes à remercier est longue. Je souhaite dédier ces mémoires, en particulier l'étude inédite, la comparaison synchronique des langues romanes, à Monsieur Joseph Herman, spécialiste du latin vulgaire et de l'histoire du français, qui fut mon directeur de thèse, qui a commencé à m'instruire sur le raisonnement scientifique et à qui je n'ai plus l'occasion de demander des conseils et de montrer mes travaux.

1. La diversité des langues dans la linguistique générale

1.1. La linguistique générale

En 1954, puis en 1963, Emile Benveniste¹ a mené des réflexions sur les tendances récentes en linguistique générale en soulignant l'importance du développement de la linguistique pendant les décennies précédentes. En ce moment, celui qui s'interrogerait sur notre discipline serait dans une situation encore plus difficile si l'objectif était de faire un bref survol du développement qui s'est effectué entre les années 1950 et de nos jours. Ce n'est évidemment pas mon objectif sur ces pages ; je propose de saisir un seul aspect dans les recherches en linguistique générale, à savoir la diversité des langues et les différents moyens qui permettent de les décrire et de les analyser. Dans un premier temps, je ferai un bref survol des réflexions des linguistes que nous citons comme sources dans nos travaux sur les questions fondamentales en linguistique générale. Le fil conducteur sera la question de savoir quelle est la place que les grands théoriciens assignent à l'étude de la diversité des langues. En effet, c'est dans cette perspective que s'inscrivent mes propres recherches dont je fais la synthèse ici.

Il paraîtrait banal de définir la linguistique générale comme étant la branche de la linguistique qui se pose les questions fondamentales et les plus générales sur le fonctionnement du langage. Toutefois, en essayant de mieux délimiter l'objet des réflexions, on arrive à des différenciations qui ne sont pas évidentes à première vue. Rastier (2001) suggère une éventuelle scission entre *linguistique générale* et *linguistique universelle*, dans la mesure où la première prend pour objet les langues, alors que la seconde traite du langage. Notons qu'une telle scission existe de fait dans les recherches linguistiques en Hongrie où la plus grande faculté de lettres connaît un département de Linguistique générale qui suit l'approche traditionnelle, et un autre de Linguistique théorique dont les activités s'inscrivent dans les différentes écoles des grammaires formelles.

Mais théoriquement, qu'est-ce qui pourrait justifier une telle distinction ? L'intégration dans une discipline plus large, les sciences cognitives, comme le prévoit Rastier, en est une raison. Une autre est la différence des méthodes utilisées. La linguistique générale, dans laquelle je propose d'intégrer la typologie, et la linguistique théorique – formelle ont les mêmes objectifs : découvrir comment fonctionne le langage humain, quels sont les invariants dans les langues et quelles sont les limites de la variation. La différence fondamentale entre les deux approches réside dans

¹ *Problèmes de linguistique générale*, tome I., 3-31.

l'opposition entre la méthode inductive de la typologie et la méthode hypothético-déductive des approches formelles. Toutefois, depuis la théorie des Principes et paramètres (Chomsky 1981) des approches formelles acceptent, elles aussi, l'idée de la variation paramétrique. Et d'autre part, certains de ces paramètres sont utilisés dans la description et dans la comparaison des langues. Nous en trouvons un exemple intéressant dans Shibatani et Bynon (1995) où Fukui fait une comparaison syntaxique de l'anglais et du japonais selon l'approche Principes-et-Paramètres.

1.2. Les linguistes sur la diversité des langues

La comparaison des langues, dans une perspective diachronique, a été particulièrement développée au XIX^e siècle. L'établissement des rapports génétiques et les tentatives d'une classification morphologique des langues en sont les témoignages. En parcourant l'histoire de la linguistique de ce point de vue, on observe que ce n'est qu'en 1921 avec Sapir que la diversité des langues retrouve sa place dans les réflexions sur le langage. Saussure est considéré comme le fondateur de la linguistique en tant que science. La question que l'on peut se poser est de savoir si dans les écrits de Saussure, aux endroits où il propose d'élaborer l'objet de la linguistique, il désigne une place à la diversité des langues. Il paraît qu'il ne s'y réfère pas explicitement quand il oppose la langue et la parole comme parties constitutives du langage. Toutefois, au moment où il analyse un phénomène linguistique concret, il propose une comparaison ou une confrontation entre deux langues. C'est le cas lorsqu'il donne un exemple² pour illustrer la notion de la valeur des signes sous forme d'une comparaison entre *mouton* du français et *mutton* de l'anglais, ce dernier s'opposant à *sheep*. De même, on pourrait citer l'exemple de *mieten* et *vermieten* en allemand face à *louer* en français. Même si cela n'apparaît pas explicitement, l'observation des faits de langue passe nécessairement par l'observation simultanée de différentes langues.

Il n'est peut-être pas exagéré d'attribuer à Sapir la relance des réflexions sur la comparaison des langues. Nous ne développerons pas en détail l'importance de la description des langues amérindiennes, il la souligne en renvoyant aux travaux de Bloomfield sur l'algonquin et aux siens sur l'athabasque. Nous notons seulement que l'analyse et la description des concepts utilisés dans ces langues mènent à relancer la réflexion sur les rapports qui relient la langue et la pensée, réflexions dont

² CLG 2^e partie, chap. 4, pp. 158-162.

nous connaissons l'importance dans les différentes approches cognitives récentes.

Le fait que la typologie morphologique proposée par Sapir (1921, 142-143) n'a pas trouvé de suite directe n'en diminue pas les valeurs. La reconnaissance du caractère souple du type morphologique et la clarification des deux paramètres, celui du degré de synthèse (mesuré par le nombre de morphèmes par mot) et le degré de fusion entre les morphèmes dans les cas où il y en a plus d'un par mot a permis de développer le concept de type linguistique, tel qu'il apparaît plus tard chez Greenberg (1954).

Mais Greenberg n'est pas le seul, à cette époque-là, à défendre l'importance des études comparatives et à promouvoir les études typologiques. Jakobson (1958), en mentionnant l'échec de la tentative de Friedrich Schlegel de créer une typologie générale des langues, trouve « naturel que Sapir, qui fut un des promoteurs de la linguistique descriptive, se soit fait l'avocat des recherches sur les types de structures linguistiques ». Pour ce qui est du rapport entre description de systèmes particuliers et comparaisons, Jakobson considère que les deux tâches, à savoir la description de systèmes isolés et l'élaboration de taxinomies s'impliquent mutuellement³. Au même moment, Jakobson se déclare être en accord avec Greenberg sur l'importance du développement d'une typologie phonologique en élaborant en même temps une terminologie plus adéquate. Ceci est une nécessité dans toute approche typologique. Mais l'une des observations les plus importantes décrivant le rapport entre les faits de langue et les lois universelles, notamment les observations qui permettent de tirer des conclusions sur la méthodologie des études typologiques est, à mon sens, la suivante⁴ :

« Le moment est venu de s'attaquer à la question des lois universelles du langage, en particulier des lois phonologiques. Même la découverte, dans une langue lointaine, nouvellement décrite, de particularités qui contredisent ces lois ne dévaluerait pas les généralisations basées sur l'étude antérieure d'un nombre imposant de langues. La régularité observée deviendrait une régularité approximative, une règle d'une haute probabilité statistique ».

Du point de vue du travail du typologue cela signifie qu'à la lumière de nouvelles données « découvertes » dans les langues certaines généralisations doivent être reformulées. Et c'est pour la même raison qu'il faut défendre la pertinence de la description de langues individuelles en

³ C'est l'une des raisons pour lesquelles la présentation de mes travaux est articulée également en descriptions et comparaisons.

⁴ Jakobson p. 72.

typologie, ainsi que la comparaison contrastive de deux ou d'un petit nombre de langues. Un tel travail est évidemment bien différent de celui d'un autre typologue qui travaillerait à l'aide de base de données informatisées qui permettent d'observer des centaines de langues à la fois.

Pour revenir aux idées des linguistes sur l'étude de la diversité des langues, nous nous arrêtons à celles d'André Martinet dont les travaux s'inscrivent exactement dans ce que nous entendons par linguistique générale. Sans parler explicitement de la diversité, il en tient compte dans toutes ses réflexions, à commencer par sa définition de la langue⁵ :

« Une langue est un instrument de communication selon lequel l'expérience humaine s'analyse, différemment dans chaque communauté, en unités douées d'un contenu sémantique et d'une expression phonique [...] ».

L'autre géant français de la réflexion linguistique, Emile Benveniste est le plus souvent cité à propos de l'énonciation ou certains autres problèmes concrets sur lesquels il a travaillé. Ici, je souhaite me référer à ses écrits, mentionnés au début de ce chapitre, sur le développement des disciplines linguistiques depuis le XIX^e siècle jusqu'aux années 1950. Il explique en effet très clairement l'importance qu'il faut attribuer à l'observation des langues les plus diversifiées.

Benveniste voit l'évolution des tendances dans l'abandon de certains problèmes, par exemple le refus⁶ « d'ériger en propriétés universelles du langage les particularités d'une langue ou d'un type de langues ». Il constate que « [l]oin de constituer une norme, le type indo-européen apparaît plutôt exceptionnel ». C'est que l'horizon des linguistes s'est élargi « par l'analyse des langues 'primitives' dans lesquelles se révèle une organisation hautement différenciée et systématique ». D'où, plus loin, sa conclusion : « Il faut attacher une grande importance à cette expérience toujours plus large des variétés linguistiques du monde ».

On pourrait s'arrêter là, ne pouvant aller plus loin dans la programmation des recherches linguistiques. Il est toutefois intéressant de jeter un coup d'œil sur deux ouvrages importants qui nous ont guidés dans nos études et que nous ne cessons de citer à l'intention des nos étudiants : il s'agit des manuels de Robins (1964) et de Lyons (1968).

Robins (1964, 15) précise le propos de la linguistique générale :

⁵ Martinet 1970, 20.

⁶ Benveniste 1966, I. 5.

« [...] le langage humain en tant que composante universelle du comportement humain et des facultés humaines. [...]. Il est bien évident qu'il n'existe pas de 'langage universel', objet spécifique de la linguistique générale, en dehors des nombreuses langues parlées dans le monde ».

Notons que c'est le premier manuel qui présente certaines manières de comparer les langues (chap. 8). Il évoque la comparaison typologique dans les domaines de la phonologie, de la phonétique et de la sémantique ainsi que dans la grammaire, avec un intérêt particulier à ce qu'il appelle « typologie grammaticale », à savoir la répartition des langues selon « isolante », « agglutinante », etc.

L'ouvrage de Lyons, *Introduction to theoretical linguistics* (traduit en français par *Linguistique générale. Introduction à la linguistique théorique*) ne développe pas explicitement le problème de la diversité, mais, comme le fait entre autres Martinet, chacune des analyses s'appuie sur plusieurs langues, celles qui sont les plus pertinentes du point de vue du phénomène en question. Il suffit de mentionner l'analyse de Lyons sur les systèmes casuels, citée jusqu'à nos jours. Toutefois, c'est pour une autre raison que nous avons cité son ouvrage original. En effet, le titre ne suggère pas, comme le fait la traduction en français, de dissocier linguistique générale et linguistique théorique.

2. Sur la comparaison des langues

2.1. Les différents types de comparaisons

La comparaison scientifique des langues peut être subdivisée en deux types de travaux⁷, avec des méthodes et des objectifs différents. On peut distinguer ainsi deux disciplines :

1. la linguistique comparative et historique et
2. la comparaison typologique.

Dans la première approche, les recherches, par définition « intragénétiques » aboutissent à l'établissement de familles linguistiques, alors que dans la deuxième on aboutit à des types ou schémas translinguistiques. Ici, il faut préciser que le terme « translinguistique » correspond au terme anglais *cross-linguistic* et se réfère au fait que les approches sont censées représenter toutes les langues du monde, à travers un échantillon représentatif. En ce sens, le terme est utilisé comme synonyme de « typologique ».

Dans les approches typologiques, lorsqu'on travaille sur un certain nombre de langues, on peut faire une distinction en fonction du rapport qui existe entre les langues qui seront étudiées. Ainsi, dans la synthèse que je propose, une distinction sera faite entre

1. comparaisons intragénétiques et
2. comparaisons intergénétiques.

Au sein des comparaisons typologiques on peut encore faire une subdivision, du point de vue méthodologique, entre approches typologiques proprement dites et études contrastives. La différence entre les deux se trouve dans l'envergure des comparaisons. En effet, la typologie, à travers d'échantillons représentatifs, s'intéresse en principe à toutes les langues du monde et vise une description globale, afin d'établir ce que c'est qu'une langue humaine, ce qui est possible et ce qui est nécessaire pour qu'un système de communication soit considéré comme une langue humaine et enfin quelles sont les limites de la variation entre les langues. En tant qu'une définition possible de la typologie Croft (1995) propose que c'est « une approche à la théorisation linguistique ou plus précisément une méthodologie de l'analyse linguistique ». Il me semble qu'une précision rendrait cette définition plus efficace. En disant que c'est une approche à la théorisation *moyennant une méthodologie spécifique* il serait plus facile d'opposer, comme on a l'habitude de le faire, les grammaires génératives –

⁷ Robins 1973, 272.

formelles et les approches typologiques – fonctionnelles. La première caractéristique de cette méthodologie spécifique consiste à tenir compte de l'ensemble des langues du monde. Une autre caractéristique est la recherche constante des possibilités d'identification de deux phénomènes grammaticaux dans deux ou de plusieurs langues différentes. L'une des plus grandes difficultés en typologie reste la définition des concepts grammaticaux.

Au sein des approches translinguistiques, les approches contrastives se limitent à la comparaison de deux ou de quelques langues dont elles étudient les micro-systèmes, selon une approche qui prend comme point de départ le concept pour arriver à l'identification du signe auquel il appartient. C'est par conséquent une approche onomasiologique qui semble être le plus approprié pour les travaux comparatifs, même si elle ne peut être exclusive. A certains moments on est obligé de prendre un signe pour point de départ (comme nous le faisons au sujet des articles, voir 8.4.).

Les distinctions qui viennent d'être proposées permettent de placer la plupart de mes recherches dans un cadre plus large. Toutefois, certaines de mes recherches portent sur la description d'une langue, le hongrois.

2.2. La terminologie

Lorsqu'on travaille sur la description d'une langue et qu'on publie ses résultats dans une autre langue, on est souvent confronté aux difficultés que peuvent poser les différences dans la terminologie. Les difficultés sont évidemment encore plus grandes lorsqu'on travaille en linguistique contrastive ou en typologie et que chacune des langues est décrite avec une terminologie qui suit les traditions de la description grammaticale de la langue donnée.

L'importance de la terminologie est développée dans l'un des travaux récents de Gilbert Lazard, linguiste particulièrement sensible à la transparence et à la définition adéquate des concepts linguistiques. Il considère que (2006, 93-94) :

« La terminologie n'est qu'une boîte à outils, c'est vrai. Mais si les outils sont mal adaptés, le travail est mal fait. Le premier soin du bon artisan est de se procurer de bons outils ou de les fabriquer lui-même ».

Il cite au même endroit Mounin qui souligne l'importance que certains des plus grands linguistes, notamment Saussure et Meillet ont attribuée à la terminologie. Selon Lazard, il faut d'abord bien définir les termes, d'où (op. cit. 94), la première nécessité d'une description claire qui

est de bien distinguer entre *signifiant* et *signifié*. Ce devoir est beaucoup plus compliqué si l'on travaille sur plusieurs langues à la fois, comme c'est le cas en linguistique contrastive ou en typologie. La première tâche consiste donc à harmoniser les terminologies et les traduire dans la langue dans laquelle on publie ses recherches, dans notre cas en français. Et au moment où on travaille dans un cadre théorique dont la langue est l'anglais, il peut y avoir des problèmes qui ne peuvent pas être résolus aisément.

Dans mes travaux, j'ai été confrontée à ce problème à plusieurs reprises. Les grammaires traditionnelles du hongrois et écrites en hongrois ont l'habitude d'utiliser une terminologie spécifique : dans la mesure du possible, tous les termes sont traduits en hongrois, les termes « internationaux » sont soigneusement évités. Quelquefois, ces termes hongrois correspondent à des concepts bien connus mais dont ils ne sont pas de simples traductions, tels que *szenvedő szerkezet* 'construction passive' (le verbe *szenved* peut être mis en rapport avec 'souffrir, subir qc'), ou *szószervezet* 'syntagme' (litt. 'construction de mots'). Dans d'autres cas le terme correspond à un phénomène connu dans la description d'autres langues mais dans la description du hongrois il s'inscrit dans un autre ensemble de concepts. A titre d'exemple, nous présenterons le problème des *jelzők* 'épithètes'. Les grammaires du hongrois en distinguent trois :

1. *minőségjelző*, litt. 'épithète de qualité', traduit par 'adjectif qualificatif', à savoir adjectif en fonction d'épithète, *belle maison*,

2. *mennyiségjelző*, litt. 'épithète de quantité', à savoir la fonction que remplit un déterminant numéral, *deux maisons* et

3. *birtokosjelző*, litt. 'épithète de possession', à savoir la fonction du possesseur dans une construction possessive, *a gyerek könyve* 'le livre de l'enfant'.

Les traductions littérales des trois termes montrent qu'en hongrois la terminologie se repose sur la fonction, en partant de la fonction épithète de l'adjectif qualificatif, alors qu'en français, il s'agit de

1. adjectif qualificatif
2. déterminant du nom
3. complément du nom.

Dans les trois cas, il s'agit des déterminants du nom dans le sens large utilisé en typologie. Les symboles y correspondant qui décrivent le groupe nominal en question sont AN, NumN, Gen N (ou l'ordre inverse). Il paraît que la terminologie du hongrois soit plus transparente dans la mesure où elle se réfère à chaque fois au fait que ce sont tous des déterminants du

nom. Lors d'un colloque organisé par le CIEH⁸, j'ai étudié l'emploi du terme *determináns* 'déterminant' qui paraît relativement récent dans la description du hongrois. Le problème ne vient pas des signifiants qui sont transparents, ils peuvent être facilement mis en rapport les uns aux autres en français et en anglais. Il s'agit d'un vrai problème des signifiés. En effet, le terme peut être utilisé dans deux sens :

1. comme « déterminants du nom », concept dont toutes les grammaires du hongrois ne tiennent pas compte et
2. dans le sens où on l'utilise en typologie, à savoir pour décrire le rapport qui existe entre *déterminant* et *déterminé* dans tout type de syntagme ou de construction endocentrique.

Il faut admettre que l'on est plus à l'aise en consultant les grammaires considérées comme « modernes », plus précisément les grammaires formelles qui sont beaucoup plus tolérantes à l'égard des termes qui sont le plus souvent directement empruntés à l'anglais. Cela permet leur utilisation aussi bien dans un texte hongrois que dans un texte français, par exemple *operátor* 'opérateur', *deiktikus* 'déictique', *expletívum* 'explétif', etc.

On pourrait avoir l'impression que les difficultés sont moins grandes et les transparences plus saillantes lors de la comparaison des langues romanes, mais ce n'est malheureusement pas le cas. Il suffit de renvoyer à la coexistence des termes « articles contractés » en français, « preposizioni articolate » en italien ou « datif et génitif de l'article » en rhéto-roman. Nous reviendrons sur ce problème plus loin, dans l'étude synchronique des langues romanes.

⁸ « Temps, espaces, langages », Colloque international organisé par le Centre Interuniversitaire des Études Hongroises, Paris 3, du 7 au 9 décembre 2006.

3. Avant les comparaisons, la description

3.1. Pourquoi le hongrois ?

J'ai choisi de travailler sur le hongrois non seulement parce que c'est ma langue maternelle, ce qui me permet de mieux évaluer les données linguistiques, mais parce qu'en lisant la littérature typologique j'ai remarqué que cette langue est assez souvent citée pour illustrer différents phénomènes, le plus souvent au sujet de la « conjugaison double » ou à propos du morphème « *-lak* ». Il y a pourtant d'autres phénomènes intéressants qui peuvent enrichir la typologie des langues.

Mais avant tout, il se pose une question pratique : pourquoi faire une nouvelle description d'une langue dont on a déjà plusieurs, notamment des grammaires non seulement en hongrois mais aussi en anglais, en français et en allemand⁹. La réponse paraît simple : parce que chaque ouvrage représente une forme spécifique de la description : il y a des grammaires traditionnelles traduites en différentes langues, mais elles risquent de poser des problèmes de terminologie. D'autres, plus modernes, suivent les grammaires formelles et posent moins de problèmes terminologiques. Toutefois, on peut observer que si les typologues évoquent un point précis de la grammaire du hongrois, le plus souvent ils ont recours aux grammaires traditionnelles de Benkő et Imre (1972) ou de Tompa (1968). Néanmoins, on trouve une description qui s'inscrit très exactement dans la perspective typologique, celle de Kenesei et al. (1998), rédigée selon le questionnaire de Comrie¹⁰, mais elle est moins souvent citée. C'est ce qui m'a amenée à publier un ouvrage qui est différent de ceux qui viennent d'être mentionnés. Il ne constitue pas une description complète de la langue, car il se concentre sur quelques aspects typologiques, ceux qui sont les mieux développés dans la littérature mais moins étudiés dans les grammaires du hongrois. L'objectif est de trouver la place du hongrois parmi les langues par l'analyse d'un phénomène donné en harmonisant la description d'une langue particulière avec celles d'autres langues. Il s'agit des domaines suivants :

- la morphologie, en particulier l'appartenance ou non à l'un des types classiques, dans le cas concret au type agglutinant,
- l'ordre des mots et la structure informationnelle de l'énoncé,
- certains processus de grammaticalisation.

⁹ Voir les références dans Sörös 2006, p.11.

¹⁰ *Lingua* 1977, vol. 42-1.

Ce sont des domaines les mieux développés ces derniers temps. En revanche, le *passif* ne fait pas partie de ces domaines, c'est plutôt le *moyen* dont l'étude a éveillé l'intérêt des linguistes. Si j'ai choisi de faire une étude détaillée sur ce sujet, c'est pour proposer une perspective d'étude qui serait plus large que les approches connues qui raisonnent en termes de transformation « actif – passif » ou de paradigme verbal. Je propose de voir dans le passif un phénomène énonciatif.

3.3. L'ordre des mots

La raison pour laquelle l'étude de l'ordre des mots en hongrois occupe une place centrale dans mes recherches est que le hongrois semble difficile à placer¹¹ dans le modèle selon la place respective de S, V et O. En fait, plusieurs observations ont été faites sur les difficultés que rencontrent ceux qui essaient d'établir l'ordre de base¹² dans une langue particulière. Pour rendre plus précise et plus applicable la définition qui se fonde sur celle proposée par Greenberg (1963), à savoir « l'ordre respectif du sujet, du verbe et de l'objet dans des phrases déclaratives avec sujet et objet nominaux », différents facteurs ont été mentionnés. On a observé entre autres la basse fréquence des énoncés dans lesquels le sujet et l'objet seraient exprimés avec des nominaux. Un autre facteur est la définitude : dans certaines langues le sujet ne peut pas être indéfini, dans d'autres un objet défini occupe une place différente de celle de la place d'un objet indéfini. Pour ce qui est de l'aspect pragmatique, dès les premières recherches on a observé que beaucoup de langues permettent plusieurs ordres de S, V et O, il importe donc d'insister sur le fait que l'étude de l'ordre de base doit s'effectuer sur des énoncés pragmatiquement neutres. Cela peut se faire si l'on met de côté d'une part les phases interrogatives et impératives et d'autre part tous les énoncés emphatiques.

Dans les développements du modèle de Greenberg¹³, la place du sujet n'est pas prise en compte, puisque entre le sujet et le prédicat il n'y a pas le même rapport tête-dépendant que celui qui relie le nom tête à ses dépendants. Toutefois, ce sont ces développements qui ont suscité des réflexions sur les prévisions que peut faire la typologie. L'ordre SVO d'une langue permettrait ainsi de prévoir la place respective des têtes (qui prennent la place de V) et des dépendants qui accompagnent V. Ces prévisions

¹¹ Dans les ouvrages cités, on trouve les références concernant la divergence des opinions sur l'ordre des mots en hongrois qui va de SVO à libre, via SOV.

¹² Sörös 2004 a.

¹³ Bartsch, Vennemann, Lehmann, voir la Bibliographie de Sörös (2006).

permettent de distinguer, théoriquement, deux types de langues « VO » et « OV », mais à quelques exceptions près, les langues ne s'avèrent pas « harmoniques » (ou, en anglais, *consistent*), c'est-à-dire que les prévisions ne se réalisent que rarement.

La variation possible de l'ordre respectif de S, V et O a été examinée de deux manières. Dans un premier temps¹⁴, j'ai testé les six variations possibles avec sujet et objet définis et indéfinis. Le corpus de 54 phrases (variations de 'L'étudiant lit un article') permet de voir que les six variations sont possibles et grammaticales, mais elles ont une interprétation énonciative différente. Bien évidemment, les phrases qui commencent par V ne peuvent pas correspondre aux conditions proposées pour l'ordre de base, les phrases qui commencent par O représentent la topicalisation, par conséquent, afin d'établir l'ordre de base, pragmatiquement neutre, il fallait se concentrer sur les deux candidats potentiels, l'ordre SVO et SOV.

L'examen d'un autre corpus, construit¹⁵, cette fois, permet de voir que c'est l'ordre SVO qui peut être considéré comme dominant et neutre. Par neutre, en hongrois, on doit entendre les énoncés dont chaque constituant est prononcé avec le même accent d'intensité. En effet, comme cela sera montré plus loin, topicalisation et focalisation peuvent se faire simplement par le changement de l'accentuation. On observe la corrélation suivante entre la définitude de l'objet (et la double conjugaison en tant que contrainte en rapport avec la définitude) et de l'aspect verbal. Tout en tenant compte de ces facteurs, pour établir l'ordre des mots en hongrois, il faut également intégrer l'aspect verbal qui a des répercussions sur l'ordre des mots, en dehors du facteur le plus important, la définitude.

Après les observations, on peut affirmer que l'ordre SVO est l'*ordre de base* ou *ordre dominant*, étant donné que cet ordre est non marqué, dans la mesure où il est soumis à moins de contraintes que l'ordre SOV. Notamment, il s'emploie avec un GN objet précédé d'un article (défini ou indéfini), le verbe peut être perfectif ou imperfectif, à l'aspect accompli ou non accompli.

On peut considérer cette observation comme un nouveau résultat concernant l'ordre des mots en hongrois, dans la mesure où elle se fonde sur une analyse systématique des données. Toutefois, si l'on affirme d'une langue qu'elle a l'ordre dominant SVO, suivant les recherches typologiques sur l'ordre des mots, on est amené à supposer qu'il s'agit en même temps d'une langue VO, ou de « type VO », c'est-à-dire d'une langue dans

¹⁴ Dans Sörös 1999a.

¹⁵ Dans Sörös 1999b, 2004a et 2006, chap. III., des corpus différents à chaque fois.

laquelle les déterminants suivent les déterminés dans un syntagme ou dans toute autre construction. Est-ce bien le cas du hongrois ?

Selon les prévisions typologiques des diachroniciens (voir note 1) l'ordre de base des constituants permet de prévoir l'ordre respectif des déterminants et des déterminés. Ainsi, dans une langue VO les déterminants du nom, tels que : adjectif, construction génitive, relative, etc. suivent le nom, tandis que les déterminants du verbe, comme l'auxiliaire, le précèdent. L'étude d'une vingtaine de traits permet de vérifier cette hypothèse. Les traits les plus importants sont : l'existence de postpositions, la place de l'adjectif épithète, la relative et les déterminants par rapport au nom, la place respectif des éléments dans une construction comparative de l'adjectif, etc.

Les résultats montrent que, contrairement aux prévisions établies en fonction de l'ordre SVO, dans l'ensemble des traits examinés il n'y a que deux qui correspondent au type VO, à savoir la place de la relative (NRel) et l'ordre respectif de l'adjectif, de l'élément de comparaison et du standard (Adj – Comp – St).

Quant à la place de la relative, il y a deux explications. Soit on accepte la statistique (Dryer 1992) selon laquelle même dans les langues OV la relative a tendance à suivre le nom, soit on tient compte de la forme de la relative également. Comme pour d'autres phénomènes, la place de l'adjectif épithète par exemple, la complexité formelle du déterminant n'est pas négligeable. C'est-à-dire que si la relative est une phrase subordonnée, alors elle a tendance à suivre le nom, tandis que si la relative a la forme d'un participe, ce qui est très fréquent en hongrois (comme en allemand, entre autres), alors la relative précède le nom.

Concernant la place des éléments dans la comparaison de l'adjectif, l'ordre représentant VO est l'ordre neutre actuellement, tandis qu'en cas d'emphase, lorsque le locuteur veut mettre en relief le standard ('par rapport à/comparé à toi, je suis plus grand'), il est possible de produire un énoncé dans lequel l'ordre correspond au type OV.

La conclusion que l'on peut tirer des observations sur le rapport qui existe entre l'ordre de base et l'ordre respectif des déterminants–déterminés semble être un résultat important pour la typologie des langues. Le hongrois est en effet une langue SVO concernant l'ordre des constituants fondamentaux, mais une langue « de type OV » concernant l'ordre des déterminants qui précèdent les déterminés. Une telle conclusion offre en même temps une nouvelle perspective à la recherche. On peut en effet se poser la question de savoir si l'ordre des constituants dans la phrase et celui des constituants de syntagmes sont soumis aux mêmes influences ? On sait que l'ordre dans les syntagmes change moins vite, il y a le témoignage des

langues germaniques modernes. Mais la question qui reste ouverte est de savoir si ces deux niveaux sont sensibles dans la même mesure aux effets pragmatiques. L'analyse diachronique que nous avons réalisée et qui compare le français et le hongrois (voir 7. plus loin) suggère que les éléments constitutifs des syntagmes sont moins susceptibles d'être affectés par des effets pragmatiques. Cela est probablement vrai pour les déterminants ou les subordinées relatives, mais beaucoup moins pour les adjectifs qualificatifs ou pour les adverbes modifiant un verbe. Ce sont toutefois des impressions que l'on a en connaissance des langues romanes, mais la problématique devrait être développée en observant d'autres langues ou familles de langues également.

Les recherches sur l'ordre des mots aboutissent à des recherches dans un autre domaine, encore pas suffisamment exploité dans les études translinguistiques : la structure informationnelle de l'énoncé.

3.4. La structure informationnelle

Si pour étudier l'ordre des mots dans les langues ou dans une langue donnée on dispose de certains cadres théoriques et d'études préalables qui permettent de travailler avec un certain nombre de critères, ce n'est pas le cas pour la structure informationnelle. La question de dépasser l'ordre dit de base est pourtant extrêmement importante, puisque les langues peuvent connaître des ordres alternatifs également d'une part, et d'autre part toutes les théories n'acceptent pas l'idée d'un ordre neutre, ce qui correspondrait justement à l'ordre de base.

On observe qu'il existe des domaines de recherches translinguistiques qui n'ont pas encore donné lieu à une théorisation suffisamment élaborée pour une éventuelle modélisation. Ainsi, si le rapport qui existe entre l'ordre des mots (ou, plus précisément l'ordre des constituants dans la phrase) et la structure informationnelle de l'énoncé semble évident, ce dernier domaine ne connaît pas encore une théorie ou des théories suffisamment générales. Plusieurs pistes ont été pourtant proposées. Dès les débuts (voir Steele 1978) il a été observé que beaucoup de langues connaissent, en dehors de l'ordre dominant, un ou des ordre(s) alternatifs. Plus tard, la typologie de Li et Thompson (1981) a attiré l'attention sur le fait que toutes les langues ne peuvent pas être décrites en termes de Sujet et Prédicats, certaines se prêtent plutôt à une analyse en termes de Topique et Commentaire. Quant à un arrière-plan théorique, c'est celui de Lambrecht (1994) qui semble être le plus largement accepté. Lors du prochain congrès

de l'ALT¹⁶ un atelier sera organisé sur ce sujet. En attendant un cadre théorique plus complet, on peut toutefois commencer à travailler sur la description des langues particulières.

En ce qui concerne le hongrois, une description de la structure informationnelle est particulièrement importante, car certains affirment que le hongrois est une langue à « ordre des mots libre ». Du côté empirique, ce n'est que l'étude détaillée de l'énoncé qui permet de vérifier cette affirmation. D'un autre côté, mes travaux illustrent le fait qu'une analyse approfondie de données a aussi l'avantage de pouvoir corroborer une hypothèse. Il s'agit en effet de la proposition des grammaires formelles (É. Kiss 2002) qui prévoient une structure unique de tout énoncé hongrois. Contrairement aux propositions développées en typologie, les grammaires formelles n'identifient pas d' « énoncé neutre » dont les énoncés « emphatiques », à savoir topicalisés ou focalisés seraient des variantes. On suppose en revanche une structure informationnelle qui permet de décrire l'énoncé en termes de Topique et Prédicat, étant donné que tout constituant de la phrase hongroise peut être topique.

La première analyse que j'ai réalisée¹⁷ propose de trouver la place du hongrois dans les études translinguistiques, en examinant simplement les moyens que cette langue utilise, en les comparant à ceux qu'il est possible d'identifier dans les études typologiques. On constate qu'en hongrois il n'existe pas de morphème spécifique ni pour la thématization, ni pour la focalisation. A partir d'un corpus construit (54 énoncés, avec objet défini, indéfini et indéfini sans article) il est possible d'affirmer que les six variantes de S, V et O sont possibles et grammaticales, il faut les évaluer du point de vue de leur valeur énonciative. L'ordre de base, déjà étudié, est SVO. En revanche, les autres variantes illustrent les énoncés que nous appelons simplement emphatiques, à savoir : OSV représente la topicalisation de l'objet direct, SVO en dehors d'être l'ordre de base, c'est-à-dire avec un accent d'intensité peut aussi être considéré comme topicalisation du sujet (topique contrastif), l'ordre SOV peut correspondre aussi à un objet focalisé, en fonction des courbes d'intonation et de l'accent.

L'analyse d'un corpus attesté¹⁸, de cent phrases, permet d'identifier la forme et la fonction de ce qui peut être topique en hongrois. On peut démontrer que tout constituant peut être topique, mais, étant donné que le pronom sujet n'est exprimé qu'en forte emphase, le meilleur candidat à être topique n'est pas le pronom personnel¹⁹. On observe en même temps que

¹⁶ Association of Linguistic Typology

¹⁷ Lors du Colloque de Caen sur « La thématization dans les langues », en 1997.

¹⁸ Sörös (2004).

¹⁹ Comme le prévoit Lambrecht (1994).

l'impression de « liberté » de l'ordre des mots vient d'une part du fait que tous les constituants majeurs peuvent être topicalisés ou focalisés, et, d'autre part, du fait qu'il y a aussi des énoncés sans topique.

Un autre problème important est de savoir si la structure informationnelle s'analyse dans l'énoncé seulement ou bien au-delà de l'énoncé. En analysant un bref texte²⁰ du même recueil, on peut montrer que sans contexte il n'est pas possible d'analyser la structure informationnelle : si le topique traduit un élément du contexte qui est soit connu, soit identifiable pour les interlocuteurs, alors pour l'identifier, il faut connaître le contexte.

A travers des analyses sur corpus construit et corpus écrit attesté il est donc possible de soutenir l'hypothèse de la structure informationnelle Topique – Prédicat de la phrase hongroise. Toutefois, un type de corpus est complètement absent des recherches que j'ai menées : n'ayant pas accès à des corpus oraux transcrits, un certain nombre de phénomènes ne peuvent pas être examinés. En effet, on peut se poser la question de savoir si en hongrois il existe un troisième constituant énonciatif, appelé *mnémème* ou *rappel*²¹, en dehors du topique et du commentaire. Les corpus construits ou attestés (écrits) ne contiennent aucun élément comparable. A un énoncé français *Il est déjà parti, Paul* ne correspond aucune traduction exacte qui rendrait compte du constituant *Paul*. En effet, comme l'expression du sujet ne se réalise pas par un pronom personnel mais par la conjugaison, le constituant *Paul* ne remplit pas le rôle de « rappel » comme c'est le cas en français. Il n'est toutefois pas impossible que des textes enregistrés et transcrits puissent témoigner de l'existence de ce constituant.

3.5. L'agglutination

Une description typologique du hongrois serait incomplète sans l'étude du type morphologique. J'ai réalisé deux études sur ce sujet, Sörös 2006 qui s'inspire des modèles typologiques proposés depuis celui de Sapir (1921) et Sörös (en préparation 2008) dans une perspective morphologique. Cette approche a été récemment proposée²² avec l'objectif de décrire et de comparer le procédé morphologique appelé « agglutination » sous ses différentes manifestations dans les langues.

²⁰ Sörös (2006, chap. III.).

²¹ Respectivement par Fernandez-Vest (2005) et Perrot (1978).

²² D'abord pour une journée d'études, ensuite pour le no. 58. de la revue LINX, coordonné par Didier Bottineau, à paraître en 2008.

Dans Sörös 2006, j'adopte Greenberg 1954 qui propose une dizaine d'indices (indice de synthèse, indice d'agglutination, indice de composition, etc.) dont l'étude permet de chiffrer l'éloignement des langues les unes des autres. L'étude d'un corpus de cent mots offre les résultats suivants. L'indice de synthèse est 2.05, c'est-à-dire que le hongrois a en moyenne deux morphèmes par mot. L'indice d'agglutination est 0.84, plus exactement 83 agglutinations sur 99 jonctions. Le nombre des mots composés est faible : 1.06. On trouve 52 éléments dérivationnels, 44 éléments flexionnels, 3 préfixes et 86 suffixes. Toutefois, en soi-même un tel comptage n'a pas de valeur, tandis que si l'on le place parmi d'autres langues, on arrive à mieux caractériser la langue.

Ainsi, dans ce système, on trouve la place du hongrois *au milieu de l'échelle de synthèse, avec deux morphèmes par mot en moyenne*, alors qu'à l'autre extrémité de l'échelle, l'esquimo est proche de quatre morphèmes par mot. Le nombre des agglutinations est élevé, ce qui permet de supposer que les bases ne subissent que peu de changement lorsqu'elles sont en contact avec des affixes. Enfin, la tendance générale des langues à la suffixation semble également démontrée.

D'autres modèles plus récents²³ tentent également de mieux définir la distinction entre agglutination et flexion. Il s'agit de se concentrer non seulement sur le caractère séparatif ou cumulatif des affixes grammaticaux, mais aussi sur l'alternance que subissent les bases sous l'influence des affixes. Et c'est le point où l'efficacité des critères semble mise en cause. En effet, afin de déterminer la dominance de l'agglutination ou de la flexion en tant que technique morphologique dans une langue, il faudrait faire des analyses très précises sur chacun des sous-systèmes : verbes, noms, pronoms, etc. Aucune analyse (d'après mes connaissances) n'a été entreprise dans de telles profondeurs. Les recherches sur la technique morphologique des langues semblent offrir encore des perspectives.

C'est la raison pour laquelle une réflexion collective a été proposée. La description que je propose pour le hongrois est différente de la précédente, dans la mesure où le point de départ est un prototype d'une langue agglutinante auquel on compare le hongrois. Une langue agglutinante idéale ou prototypique aurait les traits suivants :

- séparation des morphèmes grammaticaux
- absence d'alternance des radicaux et des suffixes
- une forme correspond à une fonction et une fonction est exprimée par une seule forme.

²³ Ceux de Plank, Haspelmath ou Plungian (voir les références dans la Bibliographie).

Après l'observation et le classement des faits, on est amené à réfléchir sur la pertinence de la typologie morphologique dans la description des langues. Suite à l'échec des tentatives de classifications des comparaisons classiques, on a accepté l'idée qu'il s'agit *plus de caractérisation que de classification* et qu'il serait préférable d'élaborer des modèles qui permettrait de raisonner en un continuum entre agglutination et fusion et de trouver la place d'une langue sur une échelle. Toutefois, les auteurs de survols récents de la morphologie²⁴ sont un peu méfiants sur la valeur de ces regroupements. Dans Sörös (en préparation 2008) j'attire l'attention sur un aspect un peu oublié de la typologie, à savoir sur l'importance que représentent les typologies dites « complexes », telle que celle de Skalička, dont l'un des mérites est la possibilité de prévision. En effet, ces typologies permettent de mettre en rapport des phénomènes phonologiques ou phonétiques (l'harmonie vocalique), morphologiques (agglutination) et syntaxiques (ordre dominant SOV). C'est dans ce sens que je vois l'importance d'une caractérisation ou même étiquetage d'une langue selon sa technique morphologique dominante.

3.6. Le passif

La raison pour laquelle je propose de travailler sur le passif est que ce sujet reste un domaine ouvert pour la recherche translinguistique. Même si beaucoup de progrès ont été faits sur certains concepts y relatifs, il n'y a pas d'ouvrage de synthèse sur les différents phénomènes convergents.

Le phénomène du passif, même s'il intéresse beaucoup de linguistes et a suscité beaucoup de travaux, ne peut pas encore être décrit au sein d'une théorie globale. Je considère comme une approche synthétisante l'article de Shibatani 1985 qui m'a servi de point de départ pour mes recherches.

La raison pour laquelle j'ai entrepris l'étude du hongrois de ce point de vue est que les manuels de hongrois, surtout ceux destinés aux étrangers, affirment qu'en hongrois il n'y a pas de passif. Si l'on observe le paradigme verbal, en effet, il n'y en a pas. Mais si l'on accepte que le passif ait fondamentalement une fonction pragmatique, celle que Shibatani appelle « défocalisation de l'agent », alors on pourrait se demander pourquoi une langue ne connaîtrait pas de moyen pour exprimer cette fonction. Une autre raison est que je ne suis pas d'accord avec une conception du passif²⁵ qui a été proposée dans le cadre des grammaires formelles.

²⁴ Carstairs-McCarthy, Katamba, voir Bibliographie.

²⁵ Alberti (1998)

Dans un premier travail²⁶, pour démontrer l'absence ou la présence du passif en hongrois, la perspective contrastive a été adoptée : j'ai dépouillé un roman et deux numéros de revues françaises à la recherche de constructions passives. J'ai traduit ces constructions que j'ai ensuite essayé de classer et analyser. Les exemples français se distinguent par la présence ou l'absence de l'expression de l'agent, ainsi que par le caractère nominal ou pronominal de l'agent, la forme verbale « être+participe passé » étant identique. Dans l'établissement du corpus, j'ai également tenu compte des constructions contenant un infinitif.

Les traductions en hongrois permettent de faire correspondre la construction française à plusieurs constructions hongroises :

- le patient reste à l'accusatif et le prédicat est à la 3^{ème} personne du pluriel ;
- le patient devient sujet et le verbe entre dans la construction « verbe d'existence+gérondif en -vA » ;
- le patient devient sujet et le prédicat est un verbe moyen.

Ce sont les solutions les plus fréquentes que l'on peut utiliser dans les traductions. Il arrive également que ce soit l'ordre des constituants qui change :

- le patient est à l'accusatif et il se met en tête de phrase, représentant ainsi un ordre « marqué », OSV.

Lorsqu'en français il y a une construction passive représentée par le participe seul, sans verbe *être*, la traduction peut se faire à l'aide d'une relative ou par un participe passé utilisé comme adjectif.

Ces solutions concernent les cas de figure dans lesquels le hongrois n'a pas recours à l'expression de l'agent autrement que par le nominatif ou par l'accusatif. L'emploi de la postposition *által*, qui correspond à 'par' existe mais semble plus restreint.

Après cette première approche du passif, qui n'est qu'une simple observation et regroupement des données, j'ai fait une analyse plus développée²⁷. Dans un premier temps, j'ai parcouru le sort qu'a subi la « vraie » construction passive du hongrois, celle qui a disparu au 19^{ème} siècle. Il s'agit d'un suffixe composé du suffixe du factitif, *-tat/-tet*, suivi du suffixe *-ik* qui sert, entre autres, à détransitiver un verbe. Ainsi, on pouvait avoir une forme passive très productive, utilisable à tous les temps

²⁶ Sörös (2003).

²⁷ Sörös (2006).

et modes. A cette construction, il était également possible d'ajouter l'agent. Cette forme deviendra plus rare vers le milieu du 19^{ème} siècle, mais, à partir du 18^{ème}, elle est concurrencée par une autre, celle que la terminologie grammaticale hongroise appelle « construction prédicative verbo–adjectivale qui se compose du verbe d'existence et d'un élément comparable au gérondif.²⁸ ».

Dans le présent cadre, je n'ai pas l'intention d'insister sur ces évaluations normatives qui accompagnent toujours les analyses du passif. Il est plus intéressant de voir comment les grammaires formelles proposent de traiter ce problème. Alberti (1998) propose une analyse identique aux deux phrases suivantes :

a) *A szoba ki van adva*
 La chambre PREV est louer–GER
 'La chambre est louée'

b) *A Tisza be van fagyva*
 La Tisza PREV est geler–GER
 'La Tisza est gelée'

Alberti considère les deux cas comme la formation du passif, qui consiste en la « suppression d'un argument ». Cette explication est plausible pour la phrase (a) qui contient un verbe transitif dont un argument peut être supprimé. Toutefois, on ne peut pas l'accepter pour la phrase (b) où le verbe est intransitif : la suppression de l'argument unique donnerait comme résultat une phrase agrammaticale. Étant donné que le verbe intransitif ne peut rester sans argument, l'explication proposée par Alberti ne peut être acceptée. Toutefois, le hongrois connaît des verbes sans argument²⁹, mais ils appartiennent tous au groupe sémantique appelé « verbes météorologiques ».

Pour terminer, on peut trouver la place des solutions que présente le hongrois dans la théorie de Shibatani 1985. Nous pouvons accepter de dire que la fonction fondamentale du passif est la « défocalisation de l'agent », sans entrer dans les différentes nuances (agent inconnu, non identifiable, etc.) Or, en hongrois, actuellement, il n'y a pas *une forme verbale spécifique* qui remplirait cette fonction et toutes ses variantes, mais il y a un *ensemble*

²⁸ Le terme *gérondif* ne correspond pas entièrement à cet élément qui a en hongrois le nom « participe adverbial », c'est pour simplifier les gloses que nous empruntons le terme au français.

²⁹ Voir Sörös et Hevér en préparation.

de constructions qui le font. J'évalue ce phénomène, avec un terme peut-être trop anthropomorphe, en disant que la langue est actuellement à la recherche d'un outil qui permet de remplir cette fonction pragmatique, puisque, un par un, les outils actuellement disponibles ne le font que partiellement.

Quoiqu'il s'agisse de la description d'un phénomène grammatical dans une langue particulière, une telle approche n'est pas sans importance du point de vue typologique non plus. En effet, l'analyse est placée dans le cadre d'un prototype et les prototypes font partie des outils les plus importants en typologie. Selon Croft (1995) la notion de prototype a été utilisée pour la première fois pour le lexique, dans les recherches portant sur l'expression des couleurs. Depuis, on raisonne en prototype non seulement à propos du passif, mais à propos du sujet également³⁰.

3.7. La grammaticalisation de l'ESPACE

La problématique de la grammaticalisation s'inscrit dans plus d'un domaine. C'est une approche diachronique dont l'étude est particulièrement intéressante dans une langue dite « agglutinante ».

L'étude de l'apparition des éléments grammaticaux semble un sujet évident dans un travail qui porte sur la présentation d'une langue dans une perspective typologique. J'ai choisi l'étude d'un domaine précis, à savoir l'expression de l'espace. Le hongrois offre de ce point de vue un matériel important, puisque le concept ESPACE est représenté par quatre types d'éléments (que nous appellerons *grams*, selon la terminologie courante des études sur la grammaticalisation):

- adverbes
- postpositions
- suffixes casuels
- préverbes.

Après le rappel de quelques notions fondamentales, j'examine les deux processus de métaphore et de métonymie en donnant des exemples du hongrois. Il s'agit avant tout, au sujet de la métaphore, de l'usage que fait le hongrois des parties du corps. Le point de départ est la liste d'Ungerer et Schmid (1996,114) que l'on peut adapter au hongrois. Il y a d'une part des termes qui ont un emploi métaphorique dans les deux ou trois langues

³⁰ Voir la proposition de Keenan (1976) intégrée dans notre étude sur l'impersonnel (Sörös et Hevér en prép.).

comparées (départ en anglais, correspondants en hongrois, traduction en français), mais tous ne sont pas exploités, tandis que selon les langues, d'autres doivent être ajoutés. En plus, les différences d'emploi sont grandes, non seulement dans l'usage, mais dans le type de formation : étant donné le caractère morphologique du hongrois, la composition est beaucoup plus fréquente qu'en anglais ou en français. Les quelques exemples qui suivent ne sont pas tous ceux qui figurent dans le livre. En effet, juste pour le plaisir de la recherche, j'ai essayé de confronter ce phénomène dans les trois langues. Certains emplois se trouvent dans chacune des trois, avec le même sens, d'autres sont présents dans deux langues sur trois et encore d'autres uniquement en anglais ou en hongrois :

<i>head of government</i>	<i>kormányfő</i>	<i>chef du gouvernement</i>
<i>heart of town</i>	<i>a város szíve</i>	<i>au cœur de la ville</i>
<i>mouth of a tunnel</i>	<i>az alagút szája</i>	<i>la bouche du tunnel</i>
<i>neck of a dress</i>	<i>ingnyak</i>	<i>col de chemise</i>
<i>arm of the sea</i>		<i>bras de mer</i>
	<i>a föld gyomra</i>	<i>les entrailles de la terre</i>
<i>nose of a boat</i>	<i>hajóorr</i>	
<i>hands of a glock</i> 'les aiguilles [mains] d'une montre'		
	<i>csésze füle</i> 'anses d'une tasse [oreilles]	

Le hongrois est toujours cité dans les ouvrages qui portent sur la grammaticalisation, au sujet des suffixes casuels qui remontent au substantif *bél* 'intestin'. Il pourrait également figurer parmi les exemples ci-dessus, mais uniquement dans la colonne du hongrois, les deux autres langues n'exploitent pas ce terme de la même manière.

C'est pourtant l'élément qui présente le plus clairement le phénomène de grammaticalisation. Au début de l'analyse³¹ je présente un tableau des quatre types d'éléments véhiculant le sens ESPACE. Si on utilise le terme de « types d'éléments » c'est parce qu'il s'agit de deux catégories lexicales et de deux affixes, les propriétés et l'autonomie de ces éléments étant différentes. On observe que chaque contenu sémantique peut

³¹ Sörös 2006, Chap. VI. 3.3. Les grams spatiaux

apparaître parmi l'un des types de morphèmes, ou dans deux ou trois, mais aucun n'apparaît dans les quatre types possibles. A comparer

- adverbe *lent* 'en bas' et préverbe *le-* 'vers le bas'
- adverbe *elöl* 'devant', postposition *előtt* 'devant, avant', préverbe *elé-*, etc.

Dans l'analyse du processus de grammaticalisation, la première étape consiste à identifier les concepts sources. Dans notre cas, ils peuvent être :

- des parties du corps, notamment les noms qui désignent 'intestin', 'dos', 'œil' ou 'poitrine' ;
- des noms relationnels, comme 'partie avant de', 'partie inférieur de', etc. ;
- des notions spatiales abstraites, 'proximité' ou 'surface'.

C'est l'évolution de la source *bel*³² 'intestin' qui permet le mieux d'esquisser le processus qui s'est déroulé. La source nominale est pourvue de désinences adverbiales primaires, notamment *-n-* ou *-t* qui représentent un sens locatif général, ensuite *-á* ou *-é* qui expriment le latif, et *-l* qui correspond à l'ablatif. Ces substantifs déclinés, par exemple *bel+n* 'intestin+locatif' peuvent être considérés comme des adverbes, au cas échéant il s'agit d'un adverbe de sens 'dedans, à l'intérieur'. Ensuite, le sens primitif s'affaiblit et cet élément acquiert le sens 'à l'intérieur', et, s'ajoutant à un nom, par exemple *haz* 'maison', prend un sens locatif (plus précisément inessif), de *haz belen* 'à l'intérieur de la maison'. A ce stade il s'agit encore d'une postposition, mais ensuite elle peut se transformer en suffixe casuel, ce qui donne, avec le même substantif, la forme actuelle *házban* 'dans la maison'.

Voici la récapitulation du processus :

bel+n > *belen* > *haz belen* > *házban*.

On peut observer en même temps, sans entrer dans les détails, l'effet de l'harmonie vocalique. En effet, dans la langue actuelle ce suffixe casuel connaît l'allomorphie : *-ban/-ben*.

Un élément adverbial de ce type peut aboutir à un autre type de *gram*, à savoir un préverbe. Ici, le point de départ peut également être le

³² *bel* est l'ancienne orthographe, aujourd'hui on écrit *bél*.

nom *bel* ‘intestin’, pourvu cette fois de l’affixe –é, latif dont le résultat est : *belé*. Selon sa distribution, cet élément donne naissance d’une part à un suffixe casuel exprimant l’illatif, lorsqu’il suit un nom,

haz bele ‘dans la maison’ > *hazba* (illatif)

D’autre part, si *bele* précède un verbe, alors il devient préverbe :

bele megy > *bemegy* litt. ‘dedans va’ / ‘il entre’.

On remarque que le chemin que parcourt *bele* bifurque, le phénomène peut donc être qualifié comme *polygrammaticalisation* :

bele —> N + *bele* > suffixe casuel

bele —> *bele* + V > préverbe

Les mêmes données peuvent également être évaluées du point de vue du processus de changement conceptuel ESPACE > TEMPS/ASPECT. En effet, le transfert métaphorique ESPACE>TEMPS se manifeste en hongrois dans l’emploi temporel de certaines postpositions et suffixes casuels, p. ex. *előtt* ‘devant/avant’, *alatt* ‘sous/pendant’, etc., alors que le transfert ESPACE>ASPECT se manifeste dans certains emplois des préverbes spatiaux, p. ex. *ki-* ‘vers l’extérieur’, *el-* ‘éloignement’, etc.

4. Comparaisons intragénétiques

4.1. Les langues romanes

Les langues romanes sont en général considérées comme un groupe de langues homogène. L'objectif de mes premiers travaux était de tester cette homogénéité supposée selon les traits étudiés dans la typologie de l'ordre des mots. Il s'agit d'une application directe des traits qui figurent dans la première esquisse de la typologie de l'ordre des mots de Greenberg (1963) et dans ceux qui peuvent être considérés comme les développements, par exemple dans les théories de Vennemann (1973) ou Lehmann (1973).

Les recherches portent sur une famille génétique dans laquelle les rapports à la « langue-mère » et ceux entre les langues semblent faciliter la décision sur le nombre des langues à décrire. Mais ce n'est qu'une illusion, même dans le cas des langues romanes. Les difficultés concernant la distinction entre langue et dialecte persistent. La décision de décrire les cinq langues nationales sous leur forme standardisée est évidente : on travaille sur le français, l'espagnol, le portugais, l'italien et le roumain. Étant donné la situation du catalan, il peut être considéré comme la sixième langue standardisée. Pour d'autres, c'est le choix du linguiste qui s'impose mais on est également influencé par la disponibilité d'une des variantes, de l'un des dialectes. Ainsi, pour l'occitan on a accès en particulier au languedocien, pour le rhéto-roman au sursilvain et pour le sarde au dialecte de Nuoro³³. Mes travaux contiennent donc des données de neuf langues ou dialectes actuellement parlés.

4.1.1. Ordre des mots

En ce qui concerne l'homogénéité supposée de la famille romane, groupe, elle peut être confirmée si l'on tient compte des traits suivants :

SVO, Pr, N-Gen, N-A

Ce sont les traits sur lesquels se fonde la typologie de Greenberg et c'est le type dans lequel toutes les langues romanes sont classées selon lui.

En développant les traits, l'homogénéité semble encore quasi-absolu, concernant les traits suivants :

³³ Voir Bibliographie.

**Dem–N, Num–N, Adj–Comp–St, N–Rel, V–Adv, Adv–
Adj.**

Toutefois, si l'on inclut certaines autres traits morpho-syntaxiques, par exemple l'ordre respectif de l'objet direct et de l'objet indirect, l'ordre des constituants en cas d'objet pronominal ou d'autres, on observe que selon ces traits les langues romanes offrent une image beaucoup plus hétérogène. Plus on introduit de traits de comparaison, plus on s'aperçoit de l'hétérogénéité du groupe, ce qui, évidemment, conduit à la description des langues individuelles. Pour ce qui est de ce dernier ensemble de traits, celui qui contient entre autres l'accusatif prépositionnel, on peut dire que l'analyse perd de son homogénéité, du moment qu'on sort du cadre préétabli de la typologie de l'ordre des mots. Mais, en effet, il s'agit en même temps de vérifier l'homogénéité d'un groupe, ce qu'il n'est pas possible de faire sans tenir compte de certaines caractéristiques spécifiques à cette famille de langues.

Un trait spécifique des langues romanes semble être le fait que, contrairement à la majorité des langues appartenant à d'autres familles, la place de l'adjectif épithète n'est pas fixe. Cette observation, valable pour l'italien dans son échantillon, a conduit Greenberg à formuler l'universal 19³⁴. On peut démontrer³⁵ que la plupart des langues romanes actuelles, à l'exception du roumain et le sarde qui n'utilisent cette possibilité que très rarement, un grand nombre d'adjectifs peuvent occuper les deux positions. Nous avons également pu prouver que les langues romanes ont une propriété spécifique : la double position des adjectifs (dans les langues qui l'autorisent) est exploitée sémantiquement. Si l'on compare le sens dans l'opposition *un grand homme / un homme grand*, et dans *une maison propre / sa propre maison* on se rend compte que la différence de sens qui peut être formulée en termes de concret/abstrait peut aller jusqu'à la présence de deux mots dont le sens est complètement différent. Ce trait semble être une particularité romane qui ne se retrouve pas dans d'autres familles, mais pas panromane puisqu'il n'est pas présent dans toutes les langues romanes.

Si l'on est à la recherche de l'homogénéité du groupe, il est également intéressant de trouver la place de ce groupe dans le type SVO/Pr/NGen/NA. C'est l'un des types les mieux représentés parmi les langues du monde, y appartiennent par exemple l'albanais, les langues bantou, le yorouba, le vietnamien, le thaï et d'autres. Dès l'introduction des traits tels que l'ordre respectif des démonstratifs et des numéraux, on observe que l'albanais suit ici la tendance observée dans les langues

³⁴ Il prévoit que dans les langues NA certains adjectifs peuvent s'antéposer au nom.

³⁵ Sörös (2007).

romanes, à savoir l'antéposition de ces déterminants au nom (DemN, NumN), alors que les autres langues appartenant au type ont l'ordre NDem, NNum. Cela revient à dire que ce type n'est homogène que selon les quatre traits initialement proposés. Il semble donc relativement facile d'établir d'une part l'homogénéité des langues romanes, d'autre part le rapport qui les lie à l'albanais.

4.1.2. Et où placer les créoles ?

Les recherches sur les créoles ont une place importante dans les recherches typologiques de ces derniers temps³⁶. Dans le cadre de la typologie de l'ordre des mots nous avons fait une étude comparative entre le français et un créole à base française, le haïtien³⁷. L'évaluation des traits que nous avons énumérés plus haut permet de conclure que le haïtien est plus proche d'un « type VO » que le français, une langue plus « harmonique » selon les tendances. En effet, en haïtien le déterminant suit le nom et l'objet pronominal suit le verbe. Ce travail ayant été simplement une introduction, nous avons projeté de poursuivre la comparaison typologique ultérieurement.

Pour revenir à la question de départ, elle reste ouverte : dans la présente synthèse, cette comparaison a été intégrée dans la partie « études intragénétiques » parce que l'on considère le haïtien comme un créole à base française. Mais est-ce un argument suffisant ?

4.2. Langues finno-ougriennes

Les langues que nous avons examinées dans Sörös et Hevér (en préparation), à savoir le hongrois, le finnois et le mordve erzya sont choisies selon nos connaissances et représentent la branche finno-ougrienne des langues ouraliennes. Dans cette famille, le choix des langues à étudier semble plus problématique que dans le cas des langues romanes.

Lors d'un aperçu typologique des langues ouraliennes Hajdú (Hajdú et Domokos 1980, 127) ne parle pas de la distinction langue et dialecte, mais de langues sœurs qui sont plus étroitement liées du point de vue génétique et aréal. Par exemple, pour représenter les langues balto-fenniques, il choisit le finnois standard, pour les langues samoyèdes du

³⁶ *Linguistic Typology* 2001, 5 – 2/3.

³⁷ Fattier et Sörös 1999.

nord, il choisit le nenets. Ainsi, sans être vraiment représentatif pour l'ensemble de la famille, le corpus des trois langues que nous étudions constitue une coupe intéressante, car il s'agit de trois langues relativement éloignées géographiquement mais dont la parenté est incontestable. On a choisi donc de travailler sur le finnois et sur le hongrois sous leur forme standard, ainsi que sur le mordve³⁸ (plus précisément sur la variante erzya) dont les données nous ont été fournies par un locuteur natif.

Notre travail comparatif porte sur l'impersonnel, sujet qui a été proposée lors d'un colloque organisé par l'Université d'Angers en 2006. Ce sujet a un intérêt particulier pour nous, puisque dans la description de ces langues ce terme n'est pas utilisé pour décrire le même phénomène que dans les langues indo-européennes. En effet, si dans ces dernières on définit comme constructions verbales dites « impersonnelles » les verbes qui ne s'emploient qu'aux 3èmes personnes, singulier ou pluriel, du type fr. *il pleut*, des verbes sans sujet possible, comme le passif impersonnel de l'allemand et des verbes qui ont un argument obligatoire à un cas autre que le nominatif, comme all. *mich friert*, etc., une telle définition est difficilement adaptable aux phénomènes que l'on rencontre dans les langues finno-ougriennes. Déjà, le terme « impersonnel » apparaît très rarement dans les grammaires et surtout pour décrire une forme du passif en finnois. Ainsi, pour pouvoir examiner les trois langues finno-ougriennes, nous avons proposé un critère unique, à savoir l'invariance en personne. Nous avons observé la présence ou l'absence d'arguments, sachant que dans ces langues il n'y a pas de « sujet postiche ». Pour le sujet, nous avons accepté de raisonner en prototype : agent, topique, a un référent animé et est exprimé par un cas non marqué.

Nous avons trouvé que les trois langues sont relativement homogènes selon les critères proposés. Nous avons réussi à identifier trois types de constructions avec sept sous-classes, dont une contient les verbes météorologiques qui n'acceptent aucun argument. Parmi les six autres sous-classes une autre, également de sens « météorologique » accepte un sujet, facultativement. Dans cinq sous-classes un sujet explicite est présent, mais nous ne trouvons pas de sujet prototypique. Il n'y a aucun agent et seuls les expérimentés ont un référent animé. De cette manière on peut conclure que la première propriété des verbes qui sont proches de ceux que l'on identifie comme impersonnels est de ne pas avoir de sujet prototypique. Le dernier type, exemplifié par un énoncé correspondant à *La table est mise* a suscité

³⁸ Le mordve est parlé par environ 1,5 million de locuteurs dans le bassin de la Volga. Il fait partie de la branche fennique, et est ainsi plus proche du finnois que du hongrois.

plusieurs remarques de la part des participants au colloque. En effet, pour les locuteurs d'une langue indo-européenne, un tel énoncé ne représente pas le phénomène de l'impersonnel mais celui du passif. En revanche, dans les langues que nous avons étudiées, nous fondant sur le critère de l'invariance en personne, il s'agit d'impersonnel, puisque par exemple en hongrois l'expression de *la table* n'est pas obligatoire. Dans une situation de communication donnée, l'énoncé *Terítve van* 'Est mise' ne peut être interprété autrement que *La table est mise*. Il s'agit donc d'une construction verbale sans argument explicite.

Nous pouvons dire que les réflexions menées sur l'impersonnel, dans le cadre translinguistique proposé lors du colloque devraient être poursuivies dans la direction du passif, en enrichissant les résultats par la description de langues qui n'ont pas encore été suffisamment étudiées de ce point de vue.

5. Comparaisons intergénétiq. Approches typologiques

5.1. La place de l'adjectif épithète dans les langues

L'un des développements du domaine de l'ordre des mots est relatif à la place de l'adjectif épithète. Parmi les universaux de Greenberg, le no. 19 dit que « lorsque selon la règle générale l'adjectif qualificatif suit le nom, il peut y avoir un petit nombre d'adjectifs qui le précèdent normalement, mais quand en règle générale l'adjectif précède le nom, il n'y a pas d'exception ». Dans l'échantillon de Greenberg, sur les 30 langues étudiées, ce sont l'italien et le gallois qui confirment cet universal³⁹.

J'ai travaillé sur cet universal dans deux directions. Premièrement, il semble évident de s'intéresser aux autres langues romanes, pour savoir si c'est un trait qui caractérise seulement l'italien qui figure dans l'échantillon de Greenberg ou bien si c'est un trait roman. Ensuite, on peut voir s'il y a d'autres familles de langues en dehors des langues romanes et celtiques qui ont les deux ordres possibles, et on peut également vérifier la validité de l'universal concernant les langues AN.

D'après les grammaires usuelles des neuf langues (voir la bibliographie de Sörös 1995), on constate que le roumain et le sarde ne permettent pas (ou au maximum pour deux adjectifs) l'antéposition de l'adjectif. Dans les autres langues, ce sont *grosso modo* les mêmes adjectifs qui ont ce comportement, tels que *bon, grand, mal, beau*, etc. Les grammaires sont également d'accord sur les explications du phénomène qui relèvent de la prosodie (les adjectifs brefs, monosyllabiques s'antéposent plus facilement), de la sémantique (l'antéposition est la position marquée, alors que la postposition est l'ordre neutre) ou de la sous-catégorisation (les adjectifs relationnels ne s'antéposent pas).

Un problème que je laisse ouvert dans cet article est développé dans Sörös 2007 : il s'agit des adjectifs qui, dans les langues romanes, changent de sens en fonction de leur position respective par rapport au nom qu'ils qualifient.

La suite des réflexions sur l'universal 19 concerne la prévision selon laquelle dans les langues AN il n'y a pas d'exception, la place de l'adjectif dans ces langues est fixe. Pour valider cette observation, il faut étudier la nature des adjectifs en général⁴⁰. En effet, lorsque dans une langue AN,

³⁹ Il faut faire un choix sur le singulier du terme « universaux ». C'est « universal » qui semble être le plus accepté.

⁴⁰ Sörös 2004 b.

comme l'anglais ou le russe, l'adjectif est suivi d'un complément, la construction devient plus complexe, plus longue qu'un simple adjectif et, dans ce cas, le groupe adjectival dans son ensemble sera postposé au nom. La deuxième partie de l'universal 19 ne peut donc être validée que si l'adjectif antéposé est un simple élément lexical. Je propose donc de faire une distinction entre « adjectifs » et « quasi-adjectifs » et ce selon le poids de l'élément en question. Cet aspect est développé par Dryer (1992) qui oppose « éléments lexicaux » et « éléments non-lexicaux » à savoir des syntagmes entièrement récursifs. Une telle distinction, selon laquelle les éléments non-lexicaux sont plus longs et plus complexes que les éléments lexicaux permet, dans mon sens, d'expliquer la prédominance de l'ordre NA dans les langues. En effet, s'il y a moins de langues AN, c'est que A est représenté dans ce cas par un simple élément lexical qui a tendance à s'antéposer, alors qu'un élément plus lourd, plus complexe a tendance à se postposer au nom, ce qui expliquerait la prédominance de NA.

5.2. Les adpositions

Les adpositions sont examinées dans Sörös 2007. Dans cet ouvrage, la plupart des analyses sont contrastives. La raison pour laquelle je considère celle des adpositions comme typologique et non pas contrastive est que la catégorie lexicale en question est étudiée selon un ensemble de propriétés qui se dégagent des analyses précédentes relatives à cette catégorie dans d'autres langues. Les observations que je fais ne portent pas sur toutes les langues qui utilisent des adpositions, elles pourraient donc être poursuivies dans l'avenir. Toutefois, elles tiennent compte de toutes les propriétés évoquées dans les recherches, à savoir la forme, l'origine, le gouvernement des adpositions, leur relation aux affixes casuels etc., ce qui permet de tirer un certain nombre de conclusions. Deux questions semblent rester ouvertes, à savoir leur appartenance à une catégorie majeure ou mineure et celle de leur caractère de tête ou de dépendant. Tout récemment, Creissels (2006, 242) prend position en acceptant leur appartenance aux catégories majeures. Creissels travaille sur plus de langues que moi-même et malgré cette plus grande envergure des recherches il est amené à admettre certaines faiblesses d'une telle prise de position. On doit reconnaître que les adpositions forment une classe dont la description pose beaucoup de problèmes. Si tel est le cas, ce sont peut-être les critères de classification qui devraient être révisés.

6. Comparaisons intergénétiq. Approches contrastives

Les analyses réalisées dans Sörös 2007 (*Typologie et linguistique contrastive*) sont des analyses contrastives, même si on essaie de les intégrer dans des théories typologiques. C'est un parcours de ces analyses qui permettra de faire quelques observations en termes de méthodologie de la comparaison, sans reprendre une à une les analyses effectuées.

Nous avons déjà signalé que les analyses contrastives portent sur un nombre restreint de langues, par opposition aux travaux typologiques de plus grande envergure qui portent sur des échantillons représentatifs. Les langues que l'on choisit pour des analyses contrastives sont celles que l'on connaît, à différents niveaux évidemment, celles dont on a de bonnes grammaires à sa disposition ou pour lesquelles des locuteurs natifs peuvent apporter leur contribution. Il n'y a donc pas les contraintes de représentativité qui se présentent lors d'une analyse typologique. Étant donné que les observations portent sur un micro-système ou sur un phénomène bien délimité, sur un concept précis, les langues choisies seront représentatives pour le phénomène donné. Par exemple, les articles ne sont présents que dans un nombre relativement peu élevé des langues, dont les langues européennes (du point de vue géographique) qui font le corpus de l'analyse.

6.1. Méthodologie : les articles

Si l'on réfléchit sur la méthodologie des comparaisons contrastives, la première question qui se pose est celle du corpus sur lequel on travaille. Lors de la présentation précédente de quelques-uns de mes travaux qui représentent une approche typologique, la question du matériel n'a pas été soulevée. Dans la plupart des analyses typologiques on a recours, en dehors des langues que l'on connaît, à des grammaires d'autres langues ou à des manuels ou travaux typologiques connus et cités dans la littérature typologique. Ce dernier type de données présente toutefois des inconvénients : les données venant de deuxième ou de nième main sont souvent très imprécises. C'est une observation que j'ai faite au sujet des exemples cités du hongrois dans des ouvrages typologiques. Si à présent je ne donne pas de références exactes sur quelques erreurs d'interprétation ou de fautes c'est parce que ce n'est pas l'objectif du présent parcours. Il s'agit simplement de souligner, par le contraste, l'importance des analyses

contrastives qui peuvent être plus précises car les données sont mieux contrôlées par des locuteurs natifs.

Pour revenir aux analyses concrètes, on observera que le corpus analysé au sujet de l'emploi des articles dans quelques langues européennes prend comme point de départ un ensemble de onze phrases préparées par Croft (1990, 7) qui les destine à une démonstration sur la différence entre la description d'un phénomène dans une langue donnée et une analyse comparative. Le fait qu'un corpus préparé à partir de l'anglais est insuffisant a déjà été signalé par Croft. La réflexion par laquelle nous poursuivons ici portera d'abord sur les avantages, ensuite sur un inconvénient d'une telle comparaison. Dans un travail typologique on constaterait simplement, comme cela figure aussi dans la première partie de l'analyse, que telle ou telle langue connaît ou non l'article, défini et/ou indéfini. Or, une analyse contrastive permet d'observer que dans les langues étudiées l'article n'est pas grammaticalisé au même niveau, c'est-à-dire que l'emploi est plus ou moins restreint : par exemple, en français et en italien dans neuf contextes sur onze un article est présent, tandis qu'en anglais, allemand et en roumain l'article n'est présent que dans quatre contextes sur onze. Cela revient à dire qu'il est insuffisant de constater la présence d'un signe dans une langue sans en avoir décrit l'emploi. Et quant aux emplois, en décrivant plusieurs langues, on s'aperçoit que quelques autres contextes doivent compléter le corpus, par exemple l'emploi des articles avec des noms propres, en particulier des noms géographiques ou le rôle des articles dans l'expression du distributif, etc. Toutefois, c'est davantage du côté théorique qu'une telle approche peut encore être développée. En effet, ce que nous appelons ici les différents « contextes » ou types d'emploi constitue un ensemble trop hétérogène. Certains, comme « indéfini, spécifique » ou « non comptable, générique », etc. se réfèrent à des contenus sémantiques, alors que si l'on évoque « nom propre » on observe l'emploi de l'article avec une catégorie lexicale précise, et si l'on prend un contexte intitulé « prédicat nominal » on est au niveau syntaxique. Cela veut dire que l'analyse gagnerait en profondeur si l'on la plaçait dans le cadre théorique de *la détermination*. Un tel cadre correspondrait également à l'une des exigences de toute comparaison translinguistique, à savoir ce serait une approche *onomasiologique*, alors que dans l'analyse présentée on se place dans une approche *sémasiologique* en comparant les différents signes. Ceci est particulièrement clair au moment où on évoque les articles partitifs du français et de l'italien auxquels ne correspond aucun signe dans les autres langues. Seulement, on s'aperçoit que l'hétérogénéité des grammaires qui sont à notre disposition représente un obstacle sérieux à la comparaison des langues dans une perspective théorique. Avant de s'engager dans une telle

entreprise, on a besoin de grammaires dont les concepts grammaticaux et la terminologie présentent plus d'harmonie afin qu'elles se prêtent plus facilement à la comparaison.

6.2. Ordre des mots

Au sujet des travaux typologiques, il a été mentionné que l'ordre des mots est l'un des domaines où un grand ensemble de critères est à la disposition de celui qui décrit une langue ou qui en compare plusieurs. C'est vrai pour les études contrastives également et le problème qui se pose reste également le même : l'observation de l'ordre neutre, s'il existe, ou celle de l'ordre des mots dans les syntagmes qui n'est pas susceptible de variation « pragmatique » que l'énoncé est possible et facile à réaliser. Toutefois, une comparaison approfondie de la structure informationnelle ne peut s'effectuer que sur des corpus enregistrés, de préférence dans les mêmes circonstances pour chaque langue comparée. Il semblerait que nous sommes actuellement en attente de tels corpus.

6.3. Temps et aspect

Peut-on trouver un moyen plus efficace pour comparer l'expression du temps et de l'aspect dans les langues que l'utilisation de traductions ? Malgré les éventuelles libertés que se permettent parfois les traducteurs, c'est, semble-t-il, la seule solution efficace depuis les années 1970, l'apogée de la linguistique contrastive dans son sens classique. Il s'agit en effet de textes attestés, où la compétence des traducteurs n'est jamais mise en cause, même si à certains endroits un traducteur linguiste choisirait une autre solution qu'un traducteur littéraire. L'objectif des analyses contrastives dans Sörös 2007 est, même si cela reste inavoué, d'attirer l'attention des futurs enseignants à qui s'adresse l'ouvrage et aussi celle des grammairiens à réviser les descriptions de l'emploi des « tiroirs verbaux » dans les langues qui en connaissent, comme le français, plusieurs au passé. Les locuteurs et apprenants de langues dans lesquelles il n'y en a qu'une forme unique (les hungarophones ne sont pas les seuls) continuent à avoir des difficultés sur ce point de grammaire.

7. Comparaisons en diachronie

C'est l'approche diachronique qui confère à deux études sur l'ordre des mots (Sóres et Marchello-Nizia 2005, Marchello-Nizia et Sóres en préparation) un statut distinct. En effet, il s'agit de comparaison intergénétiq ue ou tout simplement contrastive, puisqu'on compare le français et le hongrois, mais les observations portent sur un processus d'évolution, comparable dans deux familles de langues différentes.

Après avoir constaté le comportement « disharmonique » du hongrois, à savoir l'ordre SVO dans la phrase, mais l'ordre OV dans la détermination, il faut s'interroger sur les raisons qui l'expliquent. Pour ce faire, nous nous situons dans une perspective comparative – diachronique.

Un parallélisme est observable entre les langues indo-européennes et les langues finno-ougriennes concernant le fait qu'il y a un changement qui conduit de l'ordre SOV de la proto-langue à l'ordre SVO des langues modernes. Le proto-indo-européen et le proto-finno-ougrien ont été reconstruits comme des langues SOV, tandis que les langues modernes, entre autres les langues romanes, les langues germaniques, le hongrois, le finnois, l'estonien, etc. appartiennent déjà aux langues SVO.

Nous avons comparé, en français et en hongrois, d'une part le changement de l'ordre des constituants, et d'autre part les traits qui relèvent de l'ordre respectif du déterminant et du déterminé. Nous avons constaté qu'en français la plupart des traits relatifs à la place respective des déterminants et déterminés représentent le type VO, seuls cinq traits sont restés de « type OV », à savoir

Adv–Adj, Dém–N, Num–N, Titre–Nom, Nég–V.

En revanche, en hongrois la plupart des traits correspondent au « type OV », seuls deux traits présentent le « type VO » :

N–Rel, Adj–Marqueur–St.

La question est de savoir ce qui explique le fait que l'ordre des constituants a évolué dans la même direction dans les deux langues, alors que l'ordre des déterminants n'a pas changé en hongrois.

Au sujet du changement de l'ordre des constituants, nous avons observé que, par opposition à ce qu'affirment les diachroniciens comme Lehmann ou Vennemann, le facteur qui déclenche le changement n'est pas le sujet, mais l'objet direct. Ainsi, en français, c'est d'abord l'objet nominal dont la place se fixe à droite du verbe ; le sujet n'est pas exprimé au début

du développement et ce n'est qu'après la fixation de l'objet nominal que le sujet se place devant le verbe (l'ordre devient alors SVO).

En hongrois, c'est également un changement concernant l'objet qui déclenche le changement. Il s'agit du marquage de l'objet : au début, l'objet indéfini n'est pas marqué, tandis que le sujet et l'objet définis sont affectés d'une marque *-t*. Au cours du changement, cette marque perd sa valeur définie et elle est associée à l'objet, défini et indéfini. C'est donc l'apparition de ce marquage qui rend possible le mouvement de l'objet : la place préverbale (OV) reste celle de l'objet indéfini, la place postverbale (VO) deviendra celle de l'objet défini. Et, selon mes statistiques (présentées plus haut dans la présente synthèse), l'ordre dominant est devenu SVO.

En ce qui concerne le fait que l'ordre des déterminants et déterminés ne suit pas le même chemin dans les deux langues, nous avons formulé l'hypothèse suivante :

« Lorsque le changement de l'ordre des mots commence par le niveau des syntagmes, il a tendance à aboutir à un changement complet du type : « OV » > « VO ».

En revanche, lorsque le changement commence au niveau de la phrase, il n'affecte que les constituants fondamentaux ; les syntagmes peuvent rester inchangés, le changement SOV > SVO ne concerne donc pas l'ensemble du système. »

Nos analyses à travers l'histoire des deux langues soutiennent l'hypothèse selon laquelle la chronologie des changements est différente : en français, le changement commence par les syntagmes et le processus aboutit, tandis qu'en hongrois le processus commence par l'ordre des constituants mais les changements n'affectent pas les syntagmes.

On peut proposer comme hypothèse explicative qui reste à tester sur d'autres langues que les constituants fondamentaux sont soumis à des effets pragmatiques, tandis que les syntagmes sont moins sensibles à ces effets.

8. L'évolution de la morphologie des langues romanes : flexionnelle et analytique ?

8.1. Introduction

La partie du document de synthèse qui suit est une recherche récente, non publiée sur les langues romanes, dans laquelle je propose de mettre en œuvre la méthodologie préconisée dans ce qui précède sur un domaine que j'ai étudié sur le hongrois.

Les langues romanes sont en général considérées comme des langues « flexionnelles, analytiques »⁴¹. C'est ce que l'on trouve tout récemment (Fuchs–Meleuc 2003) dans un ouvrage de linguistique française destiné à un large public. Même si dans de tels manuels on utilise généralement ces « étiquettes », les propriétés de ces deux techniques ne sont pas faciles à saisir. Nous avons déjà abordé le problème de la distinction entre technique agglutinante et fusionnelle. Pour ce qui est des langues romanes, c'est en particulier le rapport entre flexionnel et analytique qui doit être étudié. En effet, on ne trouve pas autant d'analyses sur ce problème-ci que sur celui-là.

La morpho-syntaxe des langues romanes peut être analysée dans différents cadres théoriques. Tout récemment, Klausenburger (2000) a fait le choix de ceux de la Morphologie Naturelle et de la grammaticalisation. En citant Schwegler (1990), il soulève une question importante, la transition entre synthétique et analytique que nous allons reprendre dans la suite. La présente étude se propose d'enrichir les précédentes sur deux points: d'une part, on présentera quelques phénomènes qui ne sont pas traités par Klausenburger, par exemple l'amalgame des prépositions et des articles (ou, *articles contractés* avec la terminologie française), d'autre part on travaillera avec huit (ou quelquefois neuf) langues romanes, ce qui nous permettra d'avoir une vue plus large sur les données que celle des ouvrages cités qui ne travaillent qu'avec deux ou trois langues.

Pour commencer, nous jugeons utile de remonter jusqu'aux origines des théories concernant les techniques « synthétique » et « analytique », même s'il faut aller loin dans le temps. Toutefois, il faut préciser dans un premier temps que dans les approches modernes les termes « analytique » et « synthétique » sont utilisés pour distinguer entre les langues dans lesquelles il y a un morphème par mot et ceux qui en ont plus d'un. Or, dans

⁴¹ Sans vouloir critiquer dans les détails leur ouvrage, notons que Fuchs et Meleuc (2003) caractérisent comme tel le français, toutefois sans suffisamment de clarté concernant les deux techniques.

l'approche qui nous intéresse ici, les deux termes apparaissent en tant que deux techniques possibles au sein des langues flexionnelles (que nous avons proposé d'appeler *fusionnelles*). Dans un premier temps nous garderons cette terminologie « classique » et nous y reviendrons après une analyse détaillée.

La répartition des langues flexionnelles en synthétiques et analytiques remonte à Schlegel (1818) qui affirme que les langues flexionnelles se caractérisent par deux types : langues analytiques, langues synthétiques. Les langues analytiques sont celles qui « sont obligées d'utiliser un article devant le nom, un pronom personnel devant le verbe, qui recourent à des auxiliaires dans la conjugaison et qui remplacent la déclinaison par des prépositions, qui expriment la gradation de l'adjectif à l'aide d'adverbes ». Plus tard, au sujet de la traduction de la Bible il ajoute le « vrai passif » comme trait d'une langue synthétique.

Schleicher (1850) décrit l'évolution des langues comme une dégradation du caractère flexionnel en technique analytique. Sans reprendre ses considérations en termes de « dégradation » (il parle également de « simplification » des formes linguistiques), nous pouvons passer en revue les évolutions concernées qu'il mentionne à propos du processus qui s'est déroulé au cours de l'évolution à partir du latin vers les langues romanes. Il énumère les changements suivants :

- 1) Les désinences flexionnelles disparaissent souvent : les désinences casuelles sont remplacées par des prépositions.
- 2) Les marques modales et temporelles « perdues » sont remplacées par des auxiliaires.
- 3) Les verbes doivent être accompagnés de pronoms personnels, puisque les marques personnelles disparaissent également.
- 4) Un dernier trait, lié également à la disparition de la flexion du nom, est l'apparition de l'article défini à partir du démonstratif et celle de l'article indéfini à partir du numéral.

Notons également que Gabelentz (1891) ajoute : « si nous voulons classer les langues selon leur morphologie, alors il est insuffisant d'étudier uniquement la forme du mot, puisque ce n'est pas le mot qui est l'unité organique de la parole humaine, mais la phrase ». D'où l'importance qu'il attribue à l'étude de l'ordre des mots.

En fonction de ces phénomènes, nous pouvons dire que les langues romanes illustrent très bien le caractère cyclique des changements de

« technique morphologique »⁴². En fait, la flexion peut se décomposer, ce qui confère à la langue quelques traits analytiques. Mais dans certains cas, au cours de l'évolution, certains éléments qui font partie de la technique analytique, peuvent d'abord s'agglutiner et s'amalgamer ultérieurement. Ce caractère cyclique a déjà été décrit par Schleicher (op. cit.) et nous allons pouvoir le démontrer au sujet des langues romanes. Cette transition ne signifie toutefois pas un changement de tous les sous-systèmes, mais elle implique l'apparition de nouveaux éléments et le changement d'emploi d'éléments existants.

Avec l'objectif de bien distinguer les deux techniques, Schwegler (1990, 46–48) propose de définir le caractère « analytique » comme suit : autonomie sémantique, syntaxique, morphologique et phonologique des morphèmes dans une unité linguistique ; tandis que « synthétique » correspondrait à l'interdépendance sémantique, syntaxique, etc. entre unités linguistiques. Nous préférons étudier des données concrètes, mais nous sommes parfaitement d'accord avec Schwegler qui souligne que si cette distinction est quelquefois problématique, elle peut être maintenue, car elle est utile dans les descriptions syntaxiques, mais ne peut pas être utilisée pour l'ensemble de la langue. Dans cet esprit, nous allons séparer nos observations sur le système nominal et sur le système verbal.

Ces questions sont en rapport avec les recherches sur les processus de grammaticalisation. Mais dans un premier temps, je propose d'examiner la question suivante : En quelle mesure les langues romanes sont-elles flexionnelles ?

8.2. Analyse de la technique morphologique du français

Comme il a été signalé, la plupart des recherches typologiques récentes en morphologie portent sur la distinction agglutination–flexion. Pour une méthode d'analyse concrète, qui permet de voir le caractère scalaire des techniques, on peut avoir recours à la typologie statistique de Greenberg 1954.

Dans ce qui suit, je propose une analyse du français, d'après un corpus de 100 mots, selon sept indices strictement morphologiques⁴³.

⁴² Les descriptions les plus détaillées concernent le futur, dans Fleischmann 1983 et Bybee 1988.

⁴³ La même analyse a été faite pour le hongrois dans (Sörös 2006, chap. II.). Dans le livre, je signale les résultats d'une telle analyse concernant le français, sans présenter l'analyse en détail.

Corpus :

Malgré l'existence, consacrée par la critique, de successives « renaissances » sous le règne de tel ou tel roi, il serait peut-être plus recommandé de parler tout simplement de naissance pour le Moyen Age : avant le XII^e siècle, la littérature en langue vulgaire, c'est-à-dire écrite dans la ou les langues qui deviendront le français, n'existe pratiquement pas. Les XII^e et XIII^e siècles correspondent à une période d'expansion, d'enthousiasme et d'expériences qui ne se retrouvera à aucun moment par la suite avec la même intensité : tout est à inventer. (...) Les romans, en vers puis en prose explorent toute la gamme [des situations amoureuses.]

Anne Berthelot – François Cornilliat : *De l'amour courtois à l'humanisme*. Littérature. Textes et documents. Moyen Age, XVI^e siècle, Editions Nathan, Paris 1988, page 6.

Rappelons qu'il s'agit des traits suivants :

1. indice de synthèse : morphème par mot (abrégé en SYNTH)
2. indice d'agglutination : nombre d'agglutinations par jointure (AGGL)
3. indice de composition : morphèmes lexicaux par mot (LEX)
4. indice de dérivation : nombre de morphèmes dérivationnels par mot (DER)
5. indice de flexion : nombre de morphèmes flexionnels par mot (FLEX)
6. indice de préfixation : nombre de préfixes par mot (PREF)
7. indice de suffixation : nombre de suffixes par mot (SUFF)

	synth	aggl	lex	der	flex	préf	suff
malgré	1		1				
l'			1				
exist ence	2	1/1	1	1			1
consacr é e	3	2/2	1		2		2
par	1		1				
la	1		1				
critique,	1		1				
de			1				

success ive-s	3	½	1	1	1		1
re naiss ance s	4	2/3	1	2	1	1	2
sous	1		1	1			
le	1		1				
règne	1		1				
de	1		1				
tel	1		1				
ou	1		1				
tel	1		1				
roi,	1		1				
il	1		1				
ser ait	2	0/1	1		1		1
peut-être	1		1				
plus	1		1				
recommand é	2	1/1	1		1		1
de	1		1				
par er	2	1/1	1		1		1
tout	1		1				
simple ment	2	1/1	1	1			1
de	1		1				
naiss ance	2	0/1	1	1			1
pour	1		1				
le	1		1				
moyen-age :	1		1				
avant	1		1				
le	1		1				
douz ième	2		1	1			1
siècle,	1		1				
la	1		1				
littér ature	2		1	1			1
en	1		1				
langue	1		1				
vulgaire,	1		1				
c'est-à-dire	1		1				
écri t e	3	1/2	1		2		2
dans	1		1				
la	1		1				
ou	1		1				
les	1		1				

langue s	2	1/1	1		1		1
qui	1		1				
deviend r ont	3	1/2	1		2		2
le	1		1				
français	1		1				
n'	1		1				
existe	3	1/1	1		1		1
pratique ment	3	1/1	1		1		1
pas	1		1				
les	1		1				
douz ième	2		1	1			1
et	1		1				
treiz ième	2		1	1			1
siècle s	2	1/1	1		1		1
correspond ent	2	1/1	1		1		1
A	1		1				
une	1		1				
période	1		1				
D'	1		1				
expansion	2	1/1	1	1			1
D'	1		1				
enthousia sme	2	1/1	1	1			1
et	1		1				
D'	1		1				
expérience s	2	1/1	1		1		1
qui	1		1				
ne	1		1				
se	1		1				
re trouv er a	4	3/3	1	1	2	1	2
A	1		1				
aucun	1		1				
moment	1		1				
par	1		1				
la	1		1				
suite	1		1				
avec	1		1				
la	1		1				
même	1		1				
intens ité :	2	1/1	1	1			1

tout	1		1				
est	1		1				
à	1		1				
invent er.	2	1/1	1		1		1
les	1		1				
roman s	2	1/1	1		1		1
en	1		1				
vers	1		1				
puis	1		1				
en	1		1				
prose	1		1				
explor ent	2	1/1	1		1		1
tout e	2	0/1	1		1		1
la	1		1				
	1.35	0.81	1	0.16	0.23	0.02	0.36

L'analyse appelle quelques remarques, d'une part au niveau du découpage en morphèmes, d'autre part dans l'évaluation des agglutinations. Les unités *peut-être*, *c'est-à-dire*, *Moyen Âge* sont considérés ici comme non segmentables selon les critères proposés. Les formes *écrite* et *toute* sont analysées comme des cas de non-agglutination, d'après leur forme orale, quoique leur forme écrite permettrait leur caractérisation en termes d'agglutination. En effet, la distinction écrit/oral ne figure pas parmi les préoccupations de Greenberg ; probablement, les langues qu'il étudie ne posent pas ce problème.

Pour comparaison, voici les langues analysées par Greenberg, auxquelles nous avons ajouté le hongrois et le français.

Langue	SYNTH	AGGL	COMP	DER	FLEX.	PREF	SUFF
Vietnamien	1.06	...	1.07	.00	.00	.00	.00
Français	1.35	.81	1.	.16	.23	.02	.36
Persan	1.52	.34	1.03	.10	.39	.01	.49
Anglais	1.68	.30	1.	.15	.53	.04	.64
Hongrois	2.05	.84	1.06	.52	.44	.03	.86
Souahéli	2.55	.67	1.00	.07	.80	1.16	.41
Sanskrit	2.59	.09	1.13	.62	.84	.16	1.18
Eskimo	3.72	.03	1.00	1.25	1.75	.00	2.72

Le premier indice permet de voir le caractère scalaire de la propriété en question : l'importance de l'analyse consiste à reconnaître que les écarts sont chiffrables. On observe que même le vietnamien n'obtient pas l'indice 1 et que dans une langue qui se trouve à l'autre extrémité de l'échelle, il y a en moyenne moins de quatre morphèmes.

Le deuxième indice ne peut être évalué sans être mis en rapport avec l'indice de synthèse. En effet, si l'on ne considère que l'indice de l'agglutination, le français moderne pourrait apparaître comme une langue agglutinante, alors qu'en tenant compte de l'indice mot par morphème, on voit que le français est plutôt analytique. En mettant en rapport les deux indices, nous pouvons affirmer que, dans les cas, peu nombreux, où il y a plus d'un morphème par mot (37 jointures dans les 100 mots), la jointure se fait sans alternance (30 agglutinations sur 37 jointures). Pour la comparaison : en hongrois on trouve 99 jointures dans le corpus de 100 mots, plus précisément, dans les 58 mots segmentables, et 83 agglutinations sur les 99 jointures. Il s'agit donc de proportions, et en aucun cas d'étiquetage d'une langue comme agglutinante, simplement sur la base de cet indice. Il faut également ajouter que dans cette analyse on traite ensemble la dérivation et la flexion, ce qui diminue également le nombre des changements de radicaux en français.

De même, l'indice de flexion ne peut être évalué isolément. On observe que cet indice est plus élevé en hongrois qu'en français. Cela semble dû au fait qu'en hongrois, au niveau du système nominal, à part l'expression du nombre, il y a une catégorie de plus qu'en français qui est exprimée par affixation, à savoir le cas et les deux catégories sont exprimées par l'agglutination. Etant donné l'absence de genre en hongrois, on constate qu'en français, le nombre des catégories exprimées est identique : nombre et genre, mais la technique de fusion ne permet pas de les identifier clairement.

Il serait intéressant de faire la même analyse sur toutes les langues romanes. Concernant l'espagnol, nous avons l'évaluation de Green (1988) qui dit que l'indice de synthèse en espagnol est entre 1.9 et 2.2, et ce selon la complexité du registre. Il trouve également que beaucoup de substantifs sont fléchis pour le nombre et le genre et le système verbal reste aussi flexionnel. Comme exemple il donne une forme verbale comme

se escribían
REFL écrire–IMP–SG

dont la traduction pourrait être 'ils avaient l'habitude de s'écrire mutuellement'.

En l'absence d'analyses précises, nous fondant uniquement sur nos connaissances générales, il semble que le roumain, l'italien et le portugais seraient « plus flexionnels » que le français : dans les deux premières langues c'est la flexion nominale, dans la dernière ce sont les pronoms clitiques amalgamés dont la prise en compte permettrait sans doute d'avoir des résultats bien différents de ceux du français. Mais cela ne reste qu'une hypothèse – et une perspective de recherche.

Toutefois, une telle analyse permet de placer une langue sur une échelle chiffrable, lorsque l'objectif est une comparaison translinguistique, mais elle ne permet pas encore de dire en quelle mesure la langue est flexionnelle. Pour ce faire, une analyse de tous les sous-systèmes serait nécessaire : il semblerait qu'en français par exemple le système verbal est « plus flexionnel » que le système nominal ; encore faut-il distinguer entre système nominal et pronominal. Le travail en perspective est énorme : sauf erreur de ma part, aucune analyse de telle envergure n'a encore été faite sur les langues romanes.

8.3. De synthétique en analytique

En ce moment où il y a tant de nouvelles théories explicatives des phénomènes linguistiques, retourner jusqu'aux auteurs classiques, comme cela a été fait, peut étonner. Mais dans la présente analyse l'objectif reste de tester l'homogénéité d'une famille linguistique et non pas de corroborer ou infirmer les théories. En effet, par exemple à travers l'ouvrage cité de Klausenburger (2000) on observe avec intérêt de combien de manières différentes un phénomène linguistique peut être évalué en fonction des différentes théories.

Notre cadre étant la typologie des langues, nous allons tenter d'y placer les phénomènes morphologiques également. On a observé un rapport étroit⁴⁴ entre l'ordre des mots et le type morphologique. Ainsi, si l'ordre OV a tendance à changer en VO, le changement sera souvent accompagné d'une innovation qui consiste à coder de manière périphrastique ce qui avait été codé auparavant par la flexion. C'est dans cette optique que travaille Bauer⁴⁵, mais elle souligne que le changement principal en morphologie n'a pas été la tendance vers les formes analytiques mais l'ordre des éléments. Ainsi, au sujet des formes fléchies elle observe qu'en latin les éléments comme la marque de la gradation de l'adjectif, la désinence casuelle, les

⁴⁴ Hopper et Traugott (2003,52).

⁴⁵ P. 25–26, cité par Klausenburger (2000,29).

marques personnelles etc. ont suivi les catégories modifiées, tandis qu'en français, entre autres, on a affaire à des marques de degré antéposées à l'élément modifié : il s'agit des prépositions, des auxiliaires et des pronoms personnels. C'est effectivement une évaluation possible des faits. Mais on peut faire cette évaluation d'un point de vue différent, à savoir la transition synthétique > analytique, ce qui permet également de faire des réflexions en termes de grammaticalisation.

Un problème théorique qui se pose est de tenir compte du fait que tous les changements linguistiques ne sont pas des processus de grammaticalisation.

Par *grammaticalisation* nous entendons l'ensemble de processus de changements (sémantique, phonologique, morpho-syntaxique) qui, à partir d'un élément lexical autonome, aboutit à un morphème grammatical ou, d'un élément moins grammaticalisé aboutit à un autre, plus grammaticalisé. Nous supposons que le changement est progressif et unidirectionnel. Cette conception correspond au consensus qui se dégage des ouvrages tels que Heine, Claudi et Hünemeyer 1991, Bybee, Perkins et Pagliuca 1994, Lehmann 1985/1995, Heine et Reh 1984, Heine et Kuteva 2002, Marchello-Nizia 2006) et des travaux des prédécesseurs (Meillet 1948, Kurylowicz 1965).

Désormais, nous pouvons procéder à l'examen de la nature du comportement des éléments analytiques apparus au cours de l'évolution des langues romanes. Par rapport à l'énumération de traits de Schleicher, un regroupement selon les catégories affectées nous paraît plus logique. Ainsi, les changements concernant le groupe nominal sont :

- l'apparition de l'article défini et indéfini
- le développement de nouvelles prépositions,
- marques de gradation des adjectifs.

Quant au groupe verbal, on observe

- la généralisation des pronoms personnels sujet ; le comportement de deux clitiques obliques,
- le développement des auxiliaires, y compris ceux du passif.

Nous pouvons nous apercevoir qu'il ne s'agit pas seulement d'apparition de nouveaux éléments, mais aussi de changements d'emploi ou, dans un sens plus large, de l'évolution d'un élément moins grammatical en plus grammatical.

Notons avant de commencer que nous n'allons pas insister sur la description précise de phénomènes bien connues en linguistique romane,

mais plutôt on choisira des éléments moins bien décrits ou moins bien expliqués dans la littérature.

8.4. Les articles

L'apparition d'articles définis à partir de démonstratifs (type élément moins grammatical>élément plus grammatical) est un phénomène bien connu, mais il est considéré comme un « européisme »⁴⁶. Selon Dryer (1989), il n'y a que 8 pour cent environ des langues du monde qui ont un article. Mais, parmi les langues de l'Europe géographique, il n'y a que les langues slaves, ainsi que l'estonien et le finnois qui n'en connaissent pas, mais en finnois on est en train d'en observer l'apparition.

Il est notoire que toutes les langues romanes ont développé un article défini dont la forme actuelle remonte à deux sources (*ille* et *ipse*, ce dernier étant représenté en sarde). Etant donné notre intérêt pour l'ordre des mots, nous allons nous tourner maintenant vers la place de l'article défini : il est antéposé au nom, sauf en roumain. Il semble intéressant de réexaminer cette divergence au sein de la famille, pour voir si les nouvelles théories arrivent à enrichir l'étude du phénomène.

Bourciez (1967, 247–248) signale que, pendant un certain temps en latin, *ille*+N alterne avec N+*ille*. En Occident, c'est le premier cas qui l'emporte, tandis qu'en Orient, c'est le type *homo ille* qui gagne du terrain. Selon lui, cela peut « s'expliquer par des habitudes latines, peut-être renforcées par des influences albanaises et bulgares : il s'ensuit que l'article y devient enclitique » et forme avec le nom des groupes très étroits, par exemple, au masculin :

(1) roum.

**lupu-ille*, *lupu-llu*>*lupu-lu*>*lupul* 'le loup'

Selon Klausenburger (2000,112), en dehors de l'explication tenant compte du substrat ou adstrat balkanique, on peut considérer cela comme une option pour la solution du latin. On peut également ajouter que, du point de vue de la typologie de l'ordre des mots, c'est bien la solution dite « harmonique » de la part d'une langue VO, type vers lequel tendent d'ailleurs les langues romanes. Et en dépit de ce fait, la plupart choisissent

⁴⁶ Haspelmath (2001).

l'antéposition de l'article et des autres déterminants, en particulier du démonstratif et du numéral. Ce comportement correspond donc au type OV.

Dans un cadre typologique « européen », on observe que cette division existe aussi au sein des langues germaniques⁴⁷ où l'article peut être postposé et enclitique (sauf en présence d'adjectif épithète).

Hopper et Traugott (2003, 8–9) étudient ce phénomène du point de vue du marquage de la définitude et notent qu'en danois p. ex. ces éléments remontent à un démonstratif postposé : vieux norois : **úlfr hinn* 'loup DEM'. Toutefois, à l'heure actuelle je n'ai pas trouvé d'autre explication au fait qu'il y a ces divergences au sein d'une famille de langues. Cela reste une perspective de recherche intéressante.

Dans le cadre roman, parmi les phénomènes qui montrent une certaine hétérogénéité de la famille, notons la présence d'un article dit « partitif » en français, occitan et italien. Cet élément apparaît sous la forme de « morphème zéro » (au cas où l'on accepte le terme) dans les autres langues.

8.5. Disparition de la flexion nominale et emploi de prépositions

Dans un premier temps, il faut préciser ce que nous entendons par « disparition de la flexion ». En latin, on tient compte de la flexion selon le genre, le nombre et le cas, dans les langues romanes il faut donc examiner le sort que subissent les trois catégories.

Le marquage du nombre et du genre sur le nom n'a pas disparu. Evidemment, les réalisations sont différentes selon les langues : en français il importe de distinguer le code oral et le code écrit, en espagnol, portugais, etc. le nom est marqué à l'oral et à l'écrit par le *-s* final, le nom en roumain et en italien connaît la flexion selon le nombre, combiné avec celui de la catégorie du genre. Le roumain et certains dialectes rhéto-romans ont maintenu le neutre.

Ce qui paraît beaucoup plus intéressant c'est la généralisation selon laquelle, par rapport aux langues synthétiques, dans les langues analytiques les désinences casuelles ont été remplacées par de prépositions. Ici encore, il y a des précisions à faire.

Etant donné comme point de départ le système casuel flexionnel du latin, il n'en reste des traces qu'en roumain:

⁴⁷ Hutterer (1986).

latin : nominatif, accusatif, génitif, datif, ablatif, vocatif (+ l'ancien locatif)

roumain : nominatif=accusatif, génitif=datif, vocatif

Si l'on raisonne en termes de disparition de la flexion casuelle synthétique et emploi de prépositions qui expriment la fonction syntaxique, on a quelques observations à faire. Les cas où les prépositions prennent effectivement la fonction d'une désinence casuelle sont : le très ancien locatif (*Romae*), remplacé par différentes prépositions à sens spatial, le génitif avec la préposition *de*, le datif avec la préposition *à/a* et l'accusatif dans certaines langues et sous certaines conditions.

Il s'agit du fait que l'occitan, l'espagnol et le roumain⁴⁸ utilisent une préposition (respectivement *a*, *a* et *pe*) pour marquer l'objet direct si son référent est humain :

(2) esp. *amo a Pedro*
'j'aime Pierre'

Toutefois, on ne peut absolument pas affirmer que l'existence des prépositions serait une innovation romane. Le latin était riche en prépositions⁴⁹, qui avaient la propriété de régir des cas :

(3) latin :

avec ACC : *ante, apud, ad, etc.*

avec ABL : *a, ab, cum, sine, etc.*

ACC si répond à « Où ? » directionnel et

ABL si répond à « Où ? » statique : *in, sub, subter, super.*

En roumain également, les prépositions régissent des cas :

(4) roumain

ACC : *de, din, la, pe, prin, etc.*

DAT : *mulțumită, grație, etc.*

GEN : *contra, deasupra, etc.*

Donc, à l'exception du roumain, on peut dire que dans les langues romanes modernes, les prépositions signalent certaines fonctions

⁴⁸ Pour les références et exemples, voir Sörös (1995a) ainsi que Harris et Vincent (1988).

⁴⁹ Pendant la période du latin vulgaire en tout cas, avant, on suppose, comme dans l'indo-européen, des postpositions, dont il reste quelques-unes : *causa, gratia*.

syntaxiques des constituants nominaux⁵⁰, notamment le datif, le génitif et dans quelques cas l'accusatif.

Toujours au sujet de la flexion en général, n'oublions pas une particularité du portugais, mais qui existe aussi en sarde⁵¹, à savoir l'infinitif fléchi selon la personne et le nombre :

(5) po. *dei-lhes* *livros para lerem*
donner.1SG.PA-leur livres pour lire-3PL
'je leur ai donné des livres pour lire'.

Notons qu'au niveau translinguistique le phénomène n'est pas exceptionnel, il existe aussi en hongrois, entre autres, mais parmi les langues romanes il l'est.

8.6. La gradation des adjectifs

Les langues romanes n'ont gardé la gradation synthétique que dans le cas certains adjectifs, comme p. ex.

(7)

fr. *meilleur, pire, supérieur, etc.*
occ. *melhor, mejor, major, menor, etc.*
esp. *mejor, peor, payor, menor, etc.*
po. *melhor, pior, maior, menor, superior, etc.*
it. *maggiore, minor, miglior, peggiore, etc.*
rh. *meglier, mender, pir, meglie, etc.*

En latin certains adjectifs, se terminant en *-eus*, *-ius*, *-uus* sont également mis en comparaison de manière périphrastique, à l'aide de *magis* et *minus*.

L'apparition des éléments correspondant au français *plus*, *moins* ne signifie pas de nouveaux éléments. Il est possible de dire qu'il s'agit ici de grammaticalisation du type « élément moins grammatical > élément plus grammatical ».

⁵⁰Une telle analyse a été réalisée dans Sörös (2007).

⁵¹Parkinson (1988).

8.7. Les pronoms personnels

L'apparition des pronoms personnels qui ont pour tâche d'exprimer les catégories disparues de la conjugaison est une généralisation à laquelle il faut ajouter des précisions.

Premièrement, le pronom sujet n'est obligatoire qu'en français où il y a beaucoup de neutralisations des désinences à l'oral. En revanche, en italien ou en espagnol la « richesse » de la conjugaison ne rend pas nécessaire l'emploi du pronom sujet.

Ce qui est encore plus intéressant pour nous ici, c'est le sort que subissent les pronoms compléments. En effet, la présence simultanée d'un pronom objet direct et d'un pronom objet indirect⁵² se présente différemment selon les langues et selon l'environnement syntaxique. Comme tendance générale on observe que les clitiques précèdent les formes finies de l'indicatif et ils suivent les formes non-finies (infinitif, gérondif), ainsi que l'impératif.

En français, à l'indicatif, les deux clitiques se présentent séparément, antéposés :

(8) fr. *(je) te le donne*
(nous) les leur offrons
De même en espagnol :

(9) esp. *te lo digo*
'(je) te le dis'

En roumain, le fait que les deux pronoms sont liés est signalé dans l'orthographe par un trait d'union :

(10) roum.
le-o dă
leur-la donne
'(il) la leur donne

v-o dă
vous-la donne
'(il) vous la donne'

⁵² Nous ne traitons pas ici le problème de la place respectif des deux pronoms, voir Sörös (1995).

En revanche, en italien, le pronom objet indirect *gli* ‘lui’ et le pronom objet direct qui le suit s’amalgament :

(11) it.
glielo darò
lui-le donnerai
‘je le lui donnerai’

gliene parlerò
lui-en parlerai
‘je lui en parlerai’

En portugais, les pronoms obliques sont postposés au verbe et subissent une fusion, et ce à toutes les personnes et tous les nombres, par ex. :

(12) po. *mos*
1SG.DAT.3PL.ACC
‘me les’

lha
3SG.DAT.3SG.ACC
‘lui la’

En catalan⁵³, le paradigme est aussi riche et présente une autre particularité. Dans la plupart des variétés, les formes de *li* (3SG objet indirect) suivies d’une autre forme clitique symbolisée par X sont remplacées par X+i ou X+l s’il n’y a pas d’autre consonne. Par exemple :

(13) cat. *li+les* > *els hi* [əlzi] [əlzi]
3SG.Oi+3PL.Od
‘les lui’

Dans ces exemples, les gloses montrent bien le caractère fusionnel de ces éléments. Il s’agit donc d’un développement analytique>fusionnel, même si le phénomène ne caractérise pas toutes les langues ou pas toutes de la même manière.

⁵³ Parkinson (1988).

Après un impératif, un gérondif ou un infinitif, les pronoms ont tendance à se postposer au verbe. Dans les langues dans lesquelles on a observé une certaine fusion à l'indicatif, le phénomène est identique à celui qui caractérise les autres formes verbales. Même en français, le lien est signalé à l'orthographe par un trait d'union :

(15) fr. *donne-le-moi*

(16) it. *diteglilo*
dites-lui-le
'dites-le-lui'

per dirti
pour dire-2SG
'pour te dire'

(17) esp. *diciendotelo*
disant-te-le
'en te le disant'

dímelo
dis-me-le
'dis-le-moi'

(18) rou. *dați-mi-o*
donnez-me-la
'donnez-la-moi'

(19) po. *ofereço-lhas*
offrir-lui-les
'(je) les (FEM) lui offre'

Ces quelques exemples permettent de faire certaines observations sur la transition analytique>fusionnelle. Quoiqu'inégalement réparti, le phénomène de la fusion des pronoms obliques est bien représenté et illustre le fait que les langues continuent leur chemin d'évolution individuellement. En effet, dans certaines langues les formes ne s'amalgament pas : dans (17), en espagnol une segmentation est possible, dans d'autres cas ce n'est que l'orthographe qui signale leur lien, tandis que dans d'autres les deux pronoms apparaissent sous la forme d'un élément fusionnel. C'est le portugais qui semble avoir le système le plus riche⁵⁴.

⁵⁴ Le seul exemple du sarde est celui de Jones (1988) : *narrabilu* 'dis.lui.le', opposé à *bi lu dao* '(je) le lui donne), ou *bi* 'y' est une forme supplétive pour le pronom datif *li*.

8.8. Les auxiliaires

Un cas de grammaticalisation incontestable, du type « lexème>élément grammatical » est le développement des auxiliaires à partir de verbes pleins. Dans les langues romanes nous pouvons bien observer ce phénomène à propos de plusieurs auxiliaires. Dans un premier temps, on entend par auxiliaires ceux qui participent à la formation des temps composés, notamment les descendants de *habere* et *esse*. Le processus d'évolution est bien connu, il n'est pas nécessaire de répéter ce qui a été dit. Toutefois, au sujet de l'homogénéité supposée des langues romanes, notons que sur les neuf langues étudiées, on observe une division importante :

- *habere* et *esse* sont utilisés en français, en occitan, en italien, en sarde et en rhéto-roman ;
- seul *habere* est utilisé en catalan, en espagnol, et en roumain. En portugais, c'est actuellement le verbe *ter* qui tient lieu d'auxiliaire.

Les auxiliaires qui ont participé à la formation du futur ne sont pas les mêmes non plus⁵⁵ :

(20)	
fr.	<i>habeo+inf</i>
po.	<i>habeo de +inf</i>
it. du Sud	<i>habeo ad+inf</i>
rou.	<i>volo/habeo ad+inf</i>
rh.	<i>venio ad+inf</i>
sa.	<i>habeo ad/debeo+inf.</i>

Les auxiliaires du futur proche et du passé récent sont plus homogènes dans les langues qui les utilisent, par exemple *aller* en français, *anar* en occitan, *andare a* en italien, *ir a* en espagnol, et *venir de* en français et en occitan, *venire di* en italien, respectivement.

La formation du passif analytique est également considérée comme un phénomène panroman, mais il y a quelques différences. En général, c'est une construction *esse+participe passé* qui a pris la place de la forme synthétique du passif. Mais on doit tenir compte des deux verbes d'existence en espagnol, avec des différences d'emploi. En simplifiant un

⁵⁵ Klinkenberg (1994 ,158).

peu⁵⁶, on dit que *ser* dénote l'action ou le procès, tandis que *estar* dénote le résultat. Les premières formes sont toujours utilisées dans les écrits journalistiques ou techniques, mais dans l'usage quotidien c'est la formation avec le clitique *se* qui se répand.

En italien, à part la forme avec *esse*, on utilise également l'auxiliaire *venire* pour exprimer un procès, et *andare*, avec une nuance d'obligation :

(21) it. *questo libro va restituito*
'ce livre doit être rendu'

En portugais, en dehors de *ser et estar*, on mentionne aussi⁵⁷ les auxiliaires *ficar, ir, vir* pour exprimer un passif statif.

En roumain⁵⁸, on considère que c'est sous l'influence du français, au 19^{ème} siècle, que la construction analytique est née.

Nous observons que dans certaines langues d'autres auxiliaires viennent concurrencer *esse* dans la construction analytique. Ce qui est toutefois généralisé dans toutes les langues romanes modernes, c'est le passif exprimé avec le pronom réfléchi *se* ; c'est l'un des points les plus communs du système verbal roman.

8.9. Synthèse et analyse dans le système verbal

Nous avons regroupé la formation des « tiroirs verbaux » dans neuf langues romanes dans un tableau de synthèse (sans donner les tableaux de conjugaison détaillés) qui montre la distribution des formes analytiques et synthétiques selon les tiroirs verbaux dans les neuf langues. Nous utilisons les abréviations suivantes ici, en prenant comme base la terminologie française, y compris le terme « aoriste » :

I=indicatif, S=subjonctif, C=conditionnel, prés=présent, imp=imparfait, aor=aoriste, fut=futur, pa=passé, paC=passé composé, plqp=plus-que-parfait, ant=antérieur. Sur l'axe horizontal, S=Synthétique, A=Analytique, -=forme absente de la langue.

Le tableau ne tient pas compte des formes surcomposées, qui existent par exemple en français et qui sont très importantes en rhéto-

⁵⁶ Green (1988).

⁵⁷ Parkinson (1988).

⁵⁸ Selon Mallinson (1988). Notons que Cazacu (1981) n'en tient pas du tout compte, la conjugaison passive se présente uniquement avec le pronom réfléchi *se*.

roman. Dans cette langue, il y a encore deux formes de futur, inconnues des autres langues.

	Fr	Occ	Cat	Esp	Po	It	Rou	Sa	Rh
Prés/I	S	S	S	S	S	S	S	S	S
Imp/I	S	S	S	S	S	S	S	S	S
Aor/I	S	S	S	S	S	S	S	–	–
Fut/I	S	S	S	S	S	S	A	A	A
Prés/S	S	S	S	S	S	S	S	S	S
Imp/S	S	S	S	S	S	S	–	S	S
Fut/S	–	–	–	S	S	–	–	–	A
Prés/C	S	S	S	S	S	S	A	A	S
PaC/I	A	A	A	A	A	A	A	A	A
PaAnt/I	A	A	A	A	–	A	–	–	–
PlqP/I	A	A	A	A	S/A	A	S	A	A
FutAnt/I	A	A	A	A	A	A	A	A	A
Pa/S	A	A	A	A	A	A	A	A	A
PlqP/S	A	A	A	A	A	A	–	A	A
Pa/C	A	A	A	A	A	A	A	A	A
2 Fut spécif.	–	–	–	–	–	–	–	–	AA
Nbre tiroirs	14	14	14	15	15	14	11	12	15
Analyt/tiroir	7	7	7	7	6	7	6	8	10

Une fois de plus, on observe que l'homogénéité n'est que relative dans la famille romane. Le français, l'occitan, le catalan et l'italien semblent identiques, pour le reste des langues les écarts sont plus importants.

C'est le roumain qui semble « le plus analytique », dans la mesure où plus de la moitié de ses tiroirs, qui sont moins nombreux que dans les autres langues, sont analytiques. Ce caractère est encore renforcé par le fait qu'en roumain il y a plusieurs particules qui participent à la formation des temps composés, par oppositions aux auxiliaires conjugués d'autres langues.

De telles particules sont : *să* dans la formation du subjonctif, *o* dans celle du futur et *fi* dans celle du subjonctif passé.

8.10. D'analytique en synthétique: Les articles contractés

Tout en acceptant que les prépositions ne sont pas des éléments nouveaux alors que les articles définis le sont, il est intéressant de voir un développement morphologique dont le résultat est appelé dans la terminologie française « article contracté », au sujet de *du* et *au*. Il s'agit en effet d'un morphème « amalgame » qui se compose d'une préposition et d'un article défini. Dans les grammaires traditionnelles, les prépositions sont classées parmi les « mots invariables », et les « formes contractées » figurent parmi les articles. Mais si l'on considère l'ensemble des deux morphèmes, rien ne nous autorise à mettre en avant l'un des deux éléments. La terminologie adoptée par les grammaires des langues romanes est particulièrement hétérogène : on peut rencontrer le terme « articles contractés », ou « preposizioni articolate » en italien ou « datif et génitif de l'article » pour le rhéto-roman. Nous reviendrons sur la terminologie plus loin.

En examinant de plus près le phénomène, l'homogénéité de la famille romane semble encore une fois compromise. Toujours à l'exception du roumain, où, à cause de l'article postposé cela n'est pas possible, toutes les langues romanes connaissent le phénomène. Les différences concernent
le nombre des prépositions concernées,
les articles qui sont affectés : seuls les définis ou les indéfinis aussi,
le degré de fusion des deux éléments,
certaines neutralisations.

Ainsi, en français, la fusion ne s'effectue pas au féminin, tandis que dans d'autres langues c'est bien le cas.

La présentation suivante procède selon la « richesse » des langues en « articles contractés ». Les données représentent les prépositions et les articles concernées. Dans certaines langues, les articles indéfinis sont également fléchis, dans d'autres il y a des cas spéciaux.

esp.

Pr	M.SG.
A	AI
de	Del

+ aussi avec les pronoms personnels : *conmigo*, *contigo* etc.

fr.

Pr	M. SG	M. PL
A	au	aux
De	du	des

cat.

Pr	M.SG.	M.PL.
A	al	als
De	del	dels
Per	pel	pels

occ.

Pr	M.SG.	M.PL.
A	al, au	als
De	del, dau	dels
Per	pel	Pels
Sus	sur	Suls
Jos	jol	Jols

port.

Pr	M.SG.	M.PL.	F. SG.	F.PL.
A	ao	aos	a	as
De	do	dos	da	das
Em	no	nos	na	nas
Por	pelo	pelos	pel	pelas

+ aussi avec l'article indéfini : *numa* (*em+uma*), *dum* (*de+um*), etc.

ital.

Pr	M.SG.	M.PL.	F. SG.	F.PL.
A	al allo	ai agli	alla	alle
Da	dal dallo	dai dagli	dalla	dalle
Di	del dello	dei degli	della	delle
In	nel nello	nei negli	nella	nelle
Su	sul sullo	sui sugli	sulla	sulle
Con	col collo	coi cogli	colla	colle

+ la série avec l' (M et F) : all', dall', dell', nell', sull', coll'.

rh.

Pr	M.SG.	M.PL.	F.SG.	f.pl.
A 'à'	al agl	als	alla	allas
da 'de'	dil digl	dils	dalla	dallas
cun 'avec'	cul cugl	culs	culla	cullas
en 'dans'	el egl	els	ella	ellas
sin 'sur'	sil sigl	sils	silla	sillas
per 'pour'	pil pigl	pils	pella	pellas
sper 'à côté'	spel spegl	spels	spella	spellas
sur 'au-dessus'	sul sugl	suls	sulla	sullas
tier 'chez'	tiel tiegl	tiels	tiella	tiellas

Du point de vue de l'évolution de la morphologie, on pourrait comparer le déroulement du processus dans chaque langue, comme on le fait dans les histoires du français, en décrivant le changement qui se produit lorsqu'une [l] préconsonantique se transforme en [u]. Comme on le voit des exemples ci-dessus, ce n'est pas le cas dans toutes les langues.

Mais du point de vue de la cyclicité des techniques morphologiques, on assiste à un phénomène dont on peut mesurer la rapidité : l'émergence de l'article défini, phénomène roman, s'accompagne rapidement de cette « contraction ». Le phénomène peut, à ce stade synchronique et dans plusieurs langues, être évalué comme synthétique (fusionnel), étant donné que chacun des deux éléments subit un changement et on est certainement passé par un stade agglutinant. L'observation exacte du déroulement de ce processus dans chacune des langues et la comparaison des résultats offre un bon sujet de recherche.

Au sujet de la rapidité du phénomène, on peut ajouter les exemples que présente le français où on a connu des articles contractés avec une préposition, au cours de l'histoire de la langue, mais ils ont disparus avant le 17^{ème} siècle : *el, ou, es*.

Au sujet de la terminologie, nous avons déjà constaté qu'elle est très hétérogène. En français on appelle ces éléments « articles contractés », en italien « preposizioni articolate », traduit en anglais par « inflected prepositions »⁵⁹, pour le rhéto-roman on parle du « datif et génitif de l'article », et on pourrait continuer l'énumération. La question est de savoir quel terme est plus adéquat, celui qui insiste sur le fait que ce sont des articles (français, catalan) ou celui qui les considère comme des prépositions

⁵⁹ Napoli et Nevis, in *Phonology Yearbook* 1987, 4, 1, 195-209). Je n'ai pas eu accès à l'article même, ce n'est malheureusement que la référence que je connais.

(italien) ? Du point de vue morphologique, il s'agit de la fusion de deux éléments, une préposition et un article. Lequel a mieux gardé ses propriétés catégorielles ? Celles de l'article sont les suivantes :

- est un déterminant
- s'antépose au nom
- varie selon le genre
- varie selon le nombre.

Les propriétés de la préposition sont les suivantes :

- est un mot de relation
- s'antépose au nom
- ne varie pas selon le genre ni le nombre.

Le morphème amalgame maintient la variation selon genre et nombre (sauf quelques neutralisations) ainsi que sa place devant le nom. En revanche, il semble qu'il subit un changement de catégorie lexicale, il n'est pas un déterminant mais sa fonction est plus proche du « mot de relation ». En effet, le syntagme qu'il forme avec le nom n'est pas un groupe nominal mais un groupe prépositionnel et c'est sous cette forme-là qu'une fonction de circonstant lui est attribuée dans la phrase :

Je vais [au marché] / il sort [du parc]

Vu la fonction syntaxique que l'élément amalgamé remplit dans la phrase, il paraît que c'est celle de la préposition qui domine. Si morphologiquement il y a fusion (non-segmentabilité), alors le terme qui en résulte ne peut pas refléter le fait qu'il remonte à deux éléments autonomes. Toutefois, la fonction syntaxique permet de voir dans ce type d'éléments une « préposition fléchie » plutôt qu'un article. D'après les exemples que l'on trouve dans plusieurs langues⁶⁰ les adpositions ne sont pas toujours des mots « invariables », comme le sont considérées les prépositions dans les grammaires du français. Ainsi, je proposerai plutôt de les appeler *prépositions fléchies*, à savoir selon le genre, le nombre et la définitude.

⁶⁰ Voir Sörös 2007 sur les adpositions.

8.11. Conclusions de l'analyse

Nous venons d'étudier les phénomènes suivants dans les langues romanes modernes :

1. généralisation de l'article défini
2. disparition de la déclinaison casuelle
3. apparition de prépositions
4. article contracté
5. apparition d'auxiliaires
6. généralisation du pronom sujet
7. fusion des pronoms obliques.

Les sept phénomènes sont présents dans les neuf langues, mais leur réalisation est différente, pratiquement chaque langue a des particularités. Nous en rappelons quelques-unes.

En français, le pronom sujet est obligatoire ; seuls deux articles sont contractés et uniquement au masculin.

L'occitan est proche du français selon ces traits, à l'exception du pronom sujet qui n'est pas obligatoire.

En espagnol, deux verbes d'existence se sont développés ; la langue a maintenu le futur du subjonctif ; il y a l'accusatif prépositionnel.

En portugais, il y a des formes fléchies de l'infinitif ; on observe aussi la forte fusion des pronoms obliques ; maintien de certains tiroirs verbaux synthétiques.

En italien la flexion selon le nombre est très développée, au passif il n'y a pas seulement *esse* qui est utilisé ; il y a fusion entre deux clitiques si l'objet indirect est à la 3^{ème} personne du singulier.

Le roumain a maintenu une déclinaison nominale et le genre neutre ; a développé un article postposé ; a beaucoup plus de traits analytiques dans son système verbal que les autres.

Le rhéto-roman a développé des tiroirs verbaux qui n'existent pas dans les autres langues et utilise pour un plus grand nombre de prépositions la contraction avec l'article défini.

Notre première conclusion est plutôt banale, mais nous allons continuer la réflexion : même si l'on ne considère que quelques traits grammaticaux dans le processus de développement, les langues romanes semblent être moins homogènes que l'on ne le supposerait.

Toutefois, les mêmes phénomènes peuvent aussi être évalués d'un autre point de vue. S'agit-il de grammaticalisation dans le cas de chacun des développements ?

Selon ma conception, il y a grammaticalisation dans les trois cas suivants :

1. le démonstratif devient article défini ;
2. certains verbes autonomes deviennent auxiliaires ;
3. un certain nombre d'éléments appartenant à différentes catégories deviennent prépositions⁶¹.

Il se pose la question de savoir si l'on peut parler de grammaticalisation lorsqu'il y a fusion de deux éléments grammaticaux existants, à savoir préposition+article ou pronom objet direct+pronom objet indirect. La fusion de la préposition et de l'article est un cas spécial. La fusion aboutit-elle à un élément plus grammatical que ses composants ? Cela ne semble pas être le cas.

Le problème de la disparition de la déclinaison casuelle semble plus compliqué. La disparition des désinences pour des raisons phonétiques ne peut pas être évaluée comme grammaticalisation ; l'apparition des prépositions existantes qui seront chargées des mêmes fonctions que les cas ne me semble pas non plus satisfaire aux critères ; tout au plus il s'agit d'un changement d'emploi. C'est même le contraire : un élément de forme plus importante remplace un élément de forme plus simple.

De même, la généralisation de l'emploi du pronom sujet, qui ne caractérise que le français, ne peut pas être considérée comme l'apparition d'un nouvel élément grammatical. Toutefois, il est plus grammaticalisé, si l'on interprète sa présence (à l'oral en particulier) comme une marque indispensable de la personne et du nombre. Tel est le cas en français.

Quelques questions restent toutefois ouvertes dans cette analyse. Premièrement, la chronologie des changements, à savoir l'apparition des articles et la période où ils ont commencé à se « contracter » avec des prépositions pourrait être étudiée séparément dans chacune des langues. Ensuite, on pourrait continuer la réflexion au sujet de l'importance à attribuer à l'orthographe : la présence d'un trait d'union entre par exemple les pronoms clitiques signale-t-elle une fusion, observable à l'oral ?

Quant à une évaluation théorique des phénomènes qui viennent d'être étudiés, on observe que le caractère cyclique des changements

⁶¹ Voir Sörös (2007).

morphologiques peut être bien suivi au cours de l'histoire des langues romanes. Cette histoire est relativement brève. Une des perspectives de recherches ultérieures pourrait être une comparaison des dates approximatives des changements déroulés : l'apparition des articles contractés peut certainement être suivie dans les histoires respectives des langues. Ce que j'ai essayé de démontrer c'est que pendant une période relativement courte, à savoir dès le latin vulgaire jusqu'aux langues modernes certaines solutions analytiques, caractéristiques des transformations survenues entre le latin classique et le latin vulgaire, sont devenues fusionnelles. Cette observation ne nous permet toujours pas de nous prononcer sur le type morphologique des langues romanes, la précaution que nous avons proposée lors de l'étude du français moderne s'impose.

Les dernières conclusions portent sur la classification des langues romanes. Les manuels classiques (Vidos, Bourciez, Tagliavini, etc.) ont l'habitude de présenter les tentatives de classification, et ce avec des arguments pour et contre. Les traits les plus souvent impliqués dans les regroupements, à savoir le traitement du *-s* final et le sort des consonnes intervocaliques sourdes permettent peut-être de distinguer entre Romania Occidentale et Romania Orientale. Mais chacun sait que la prise en compte de l'ensemble des traits phonétiques, grammaticaux et lexicaux ne rendrait pas possible un regroupement interne des langues de cette famille dont l'histoire est pourtant bien documentée. La présente contribution, qui souligne le fait que des phénomènes panromans se développent de manières bien divergentes et aboutissent à des langues individuelles qui se ressemblent de moins en moins, va également dans ce sens.

Conclusions et perspectives

Dans ces conclusions, on ne reprendra pas les acquis des analyses grammaticales évoquées précédemment. Dans un premier temps, on se concentrera sur le côté méthodologique des comparaisons pour arriver à des conclusions du côté théorique également.

Pour ce qui est de la méthodologie des comparaisons synchroniques, la première étape consiste dans le choix du *domaine de la comparaison*. Le point de départ des analyses peut être un concept grammatical préétabli, comme celui des classes de mots (p. ex. les adpositions), ou les catégories grammaticales (passif, impersonnel). Le problème le plus important est le fait que la plupart des concepts grammaticaux ont été élaborés d'après les données fournies par les langues indo-européennes et qu'ils ne permettent pas toujours de travailler sur d'autres langues. Le choix d'un concept pour point de départ s'inscrit bien dans l'onomasiologie, approche préconisée dans les analyses comparatives. Toutefois, pour qu'un concept grammatical puisse rendre possible la description d'un grand nombre de langues (à la rigueur, en typologie, de toutes les langues), le travail du linguiste devra suivre ce qui a été suggéré par Jakobson : il importe de modifier ou de nuancer la définition préliminaire d'un concept chaque fois quand intervient une nouvelle donnée pour laquelle la définition proposée au début ne semble pas adéquate. Néanmoins, les définitions continuent de poser des difficultés. Au lieu de proposer des définitions, une autre possibilité consiste à établir un *prototype* à partir des traits caractéristiques. Il faut reconnaître que le raisonnement en prototypes, emprunté à la sémantique cognitive, est un outil important en typologie. C'est ainsi qu'on arrive à mieux travailler sur des concepts difficiles à définir, tels que « le passif » ou « le sujet » et à assurer une description adéquate de phénomènes qui ne font que graviter autour du prototype.

Notons encore que ce sont les concepts que Gilbert Lazard (1999) appelle *catégories interlangues*, ou plutôt *quasi-catégories*, à cause justement des difficultés de définition. Il semblerait que, entre autres, le domaine des « constructions impersonnelles » se prête également à une description en prototype, étant donné que ce concept repose, comme tant d'autres, sur les phénomènes représentés par les langues indo-européennes. La tentative d'analyse dans quelques langues finno-ougriennes, évoquée plus haut, illustre bien la démarche qui consiste à développer une définition au moment où les nouvelles données relevant de langues dans lesquelles le concept en question n'a pas encore été analysé le nécessitent. Dans le cas concret, il s'agit en plus de mettre en relation deux concepts problématiques, à savoir l'impersonnel et le sujet. Il nous sera possible de revenir sur cette

question après avoir étudié la synthèse des articles qui seront bientôt publiés.

Dans certains cas, on adopte malgré tout l'approche sémasiologique, en particulier lorsque le point de départ est une analyse contrastive qui insiste plus sur les différents types d'emploi d'un signe dans les langues que sur le concept qui le définit. C'est le problème que nous avons évoqué lors de l'analyse comparée des articles dans un certain nombre de langues européennes. L'étape suivante qui s'impose et qui nous reste à faire consiste à intégrer cette analyse dans celle du concept général qui est représenté dans les signes en question ; au cas échéant la « définitude ». C'est une des perspectives qu'offre l'analyse déjà effectuée. En même temps, on peut réfléchir sur le rapport entre analyse contrastive et analyse typologique. La différence fondamentale entre les deux n'est pas simplement celle du nombre des langues confrontées, ni la nature du concept en question. Plus loin, je propose une confrontation des deux types d'analyses.

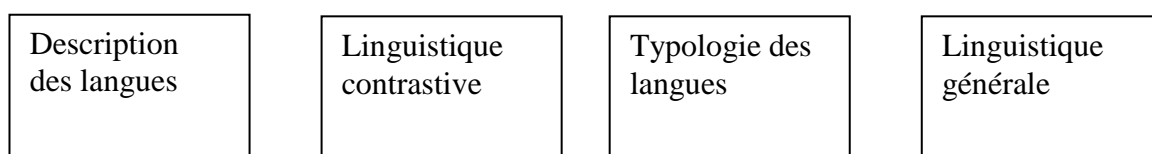
Une autre possibilité de poser un problème consiste à choisir une catégorie cognitive comme « espace » ou « possession et appartenance ». C'est le niveau le plus abstrait d'où peut partir une analyse, dans la mesure où l'on quitte le niveau des phénomènes de surface déjà lors de la délimitation de l'objet à étudier. Après avoir décrit et analysé les éléments linguistiques qui expriment le concept en question, on revient dans les explications en termes de processus cognitifs.

Le second problème méthodologique qu'il convient d'évoquer est celui des données observées. Elles sont hétérogènes, évidemment. Déjà Greenberg a précisé que son échantillon se repose sur un ensemble de langues dont certaines sont connues par le linguiste même, pour d'autres il y a des grammaires, tandis que d'autres données sont fournies par des locuteurs natifs. Comme mes travaux sont individuels ou réalisés avec un co-auteur qui apporte ses propres connaissances sur d'autres langues, la représentativité de l'échantillon ne peut être atteinte en aucune manière. Toutefois, on peut espérer pouvoir élargir les échantillons si l'on arrive à éveiller l'intérêt des jeunes chercheurs de MODYCO (qui maîtrisent la méthodologie du travail avec les bases de données informatisées) pour les analyses translinguistiques.

D'un autre point de vue se pose également la question de la/les variante/s à choisir. Quand c'est possible, l'on choisit le standard, dans d'autres cas un dialecte qui est mieux décrit que les autres et ainsi de suite. Cela donne un ensemble de données hétérogène dans un sens, mais représentatif dans la mesure où l'on s'interroge sur toutes les possibilités qu'offrent les langues, quelle que soit la variété prise en compte dans les analyses.

Si l'on accorde tant d'importance à la méthodologie c'est qu'elle permet de définir la typologie des langues en tant que discipline allant vers la linguistique générale. La typologie peut être envisagée comme une discipline décrivant le domaine intermédiaire qui se situe entre la description de langues particulières et les traits universels. Selon sa définition courante qui inclut également la recherche des universaux, elle propose d'établir les limites de la variation des structures linguistiques et de définir ce qui est nécessaire et ce qui est possible pour qu'un système de communication puisse être considéré comme une langue humaine.

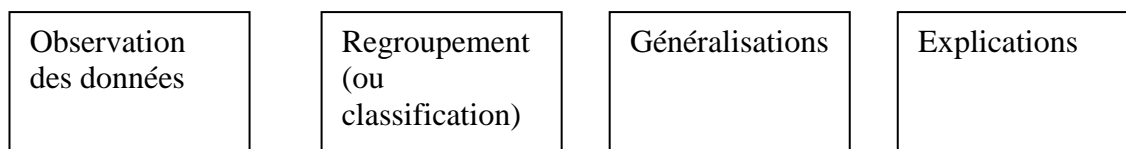
Le schéma suivant met en rapport les branches de la linguistique qui s'occupent de la description et de l'analyse des structures de superficie des langues :



Après avoir évoqué quelques problèmes méthodologiques, il convient de mettre en rapport les deux disciplines comparatives qui se situent entre la description des langues particulières et l'approche aux questions les plus générales. En comparant la linguistique contrastive et la typologie des langues, on trouve des ressemblances et des différences. Ce qui est commun, c'est évidemment l'idée de la comparaison. En outre, les analyses peuvent porter dans les deux disciplines sur les mêmes phénomènes, à savoir sur des micro-systèmes (par ex. le système vocalique des langues), ou sur une catégorie lexicale donnée (par ex. les adpositions) ou encore sur un concept grammatical ou cognitif, etc. En revanche, c'est dans la méthodologie que l'on trouve des différences. Une analyse contrastive peut travailler sur deux ou sur quelques langues, tandis qu'une analyse typologique par excellence utilise un échantillon représentatif des langues du monde. La deuxième différence importante consiste dans l'utilisation des corpus construits ou attestés dans les analyses contrastives, alors qu'en typologie on ne construit pas de corpus mais on se fie à des données relevées dans les grammaires des langues étudiées. On peut également confronter les deux approches selon leur caractère pratique ou théorique. La linguistique contrastive est considérée comme une branche de la linguistique appliquée, mais la méthode contrastive dont l'importance est soulignée ici, qui sert à décrire les phénomènes linguistiques dans une perspective comparative peut ne pas servir directement des objectifs

pratiques ou didactiques. En même temps, elle ne s'efforce pas de donner une explication théorique des phénomènes examinés. Par opposition, une analyse typologique peut s'inscrire dans une théorie partielle, comme c'est le cas des travaux sur l'ordre des mots qui essaient d'établir un ordre de base, ou bien lorsqu'on raisonne sur « tête et dépendant » dans un syntagme. La typologie se propose en même temps de trouver des explications des phénomènes, à l'aide de concepts (iconicité, économie, caractère marqué ou non marqué) et d'outils (hiérarchies, prototypes) propres à cette discipline. Toutefois, une étude typologique à grande échelle, ou plus précisément portant sur un large échantillon permet seulement l'observation de la présence ou de l'absence d'un phénomène donné, comme on l'a vu au sujet des articles. Elle ne permet pas de comparer le niveau de grammaticalisation dans les langues particulières et surtout pas les différences d'emploi. D'où l'importance représentée par la démarche inductive choisie dans mes travaux. Elle consiste à commencer l'analyse dans un nombre peu élevé de langues et à placer ensuite le résultat de la comparaison dans un cadre ou dans une théorie partielle offerte par la typologie.

Le schéma des disciplines que l'on vient de proposer doit toutefois être mis en rapport avec celui qui représente la démarche typologique, à savoir :



Le rapprochement des schémas conduit à poser la question de savoir si la typologie est une théorie linguistique ou non. Certains (Whaley 1997, 7) sont catégoriques : la typologie n'est pas une théorie grammaticale comme celle du Gouvernement et du Liage, la grammaire relationnelle ou autres qui modélisent comment fonctionne le langage. D'autres (Croft 1991, 1, 246) la considèrent comme une approche à la théorisation linguistique ou plus précisément une méthodologie de l'analyse linguistique qui permet l'accès à différentes théories, mieux que d'autres approches et elle aide la compréhension de la nature de la grammaire et du langage.

Ma proposition consisterait de dire que la typologie des langues deviendra une théorie globale au même titre que la grammaire générative au moment où elle proposera une explication de tous les phénomènes observés. Les explications proposées dès les débuts de la typologie évoquent

l'iconicité, la fréquence, le caractère marqué ou non marqué, etc. et, plus récemment les phénomènes analysés par les sciences cognitives. Le meilleur exemple en ce sens est l'étude des processus de grammaticalisation.

Pour ce qui est des perspectives qui s'offrent, il me paraît clair que, étant donné la méthodologie proposée, c'est un travail en équipe qui pourrait permettre non seulement d'élargir les domaines à étudier mais aussi d'enrichir les résultats.

ABBREVIATIONS

A(DJ)	adjectif
ABL	ablatif
ACC	accusatif
ADV	adverbe
COMP	comparaison
DAT	datif
DEM	démonstratif
FEM	féminin
GEN	génitif
GER	gérondif
IMP	imperfectum
N	nom
NEG	negation
NUM	déterminant numéral
O	objet direct
Od	objet direct
Oi	objet indirect
PA	passé
PL	pluriel
Pr(ep)	préposition
PREV	préverbe
SG	singulier
St	standard (de comparaison)
V	verbe

BIBLIOGRAPHIE

- ALBERTI Gábor, 1998, « On passivization in Hungarian », in C. de Groot – I. Kenesei (éds.) *Approaches to Hungarian*, vol. 6. Szeged, JATE : 105-121.
- ALIBERT Loïs, 1976², *Grammatica Occitana*, Montpellier, CEO.
- AR MERSER Andreo, 1983, *Précis de grammaire bretonne*, Ar Helenner.
- BADIA MARGARIT A., 1962, *Gramática catalana*, Madrid, Gredos.
- BATTAGLIA S. – PERNICONE, V., 1968, *La grammatica italiana* (2.) Torino, Loescher Editore.
- BAUER Brigitte, 1995, *The Emergence and Development of SVO Patterning in Latin and French. Diachronic and Psycholinguistic Perspectives*. Oxford, Oxford University Press.
- BOSQUE, Ignacio – DEMONTE, Violeta (ed.), 1999, *Gramática descriptiva de la lengua española*, I-II. Espasa Calpe, Madrid.
- BENKŐ Loránd – IMRE Samu, *The Hungarian Language*, The Hague Mouton, 1972,
- BENVENISTE Émile, 1966, *Problèmes de linguistique générale*, Paris, Gallimard, I-II.
- BYBEE Joan, 1985, *Morphology*, Amsterdam – Philadelphia, John Benjamins.
- BYBEE Joan – Revere PERKINS – William PAGLIUCA, 1994, *The evolution of grammar. Tense, aspect and modality in the languages of the world*, Chicago-London, The University of Chicago Press.
- CAZACU Boris et. al. , 1981, *Cours de langue roumaine*, Editura didactica si pedagogica, Bucuresti.
- CANDINAS Theo, 1982, *Romontsch sursilvan* ; Ligia Romontscha, Cuera.
- CARSTAIRS-MCCARTHY, A. (1992), *Current Morphology*, London, Routledge.
- CHOMSKY Noam, 1981, *Lectures on Government and Binding*, Dordrecht, Foris.
- CUNHA Celso – CINTRA, Luis F. L., 1987 (4^e éd.), *Nova Gramática do Português Contemporâneo*, Ed. Joao Sa da Costa, Lisboa
- COMRIE Bernard, 1989², *Language Universals and Linguistic Typology*, Basil Blackwell, Oxford.
- CREISSELS Denis, 1991, *Description des langues négro-africaines et théorie syntaxique*, Grenoble, ELLUG, Université Stendhal.
- 2006, « Suffixes casuels et postpositions en hongrois », BSLP tome CI – fasc. I. pp. 225-275.

- CROFT William, 1990, *Typology and Universals*, Cambridge, Cambridge University Press.
- 1995, « Modern Syntactic Typology », in Shibatani – Bynon, pp. 85-144.
- DEZSŐ László, 1982, *Studies in Syntactic Typology and Contrastive Grammar*, The Hague–Paris–New York, Mouton.
- DRYER Matthew S., 1992, « The Greenbergian Word Order Correlations », *Language* 68/1: 81-138.
- 1998, « Article – noun order », in Chicago Linguistic Society, 25, pp. 83-97.
- DUBOIS Jean et al. 1973, *Dictionnaire de linguistique*, Paris, Larousse.
- DUBOIS Jean – LAGANE, René, 1974, *La nouvelle grammaire du français* Paris, Larousse.
- É. KISS Katalin 2002, *The Syntax of Hungarian*, Cambridge, Cambridge Syntax Guides.
- É. KISS Katalin – Ferenc KIEFER– Péter SIPTAR, 1998, *Új magyar nyelvtan* [Nouvelle grammaire du hongrois], Budapest, Osiris.
- FATTIER Dominique – Anna SŐRES, 1999, « L'ordre des mots en créole haïtien : une introduction », *Etudes Créoles*, XXII, 2 : 48-56.
- FERNANDEZ–VEST M.M. Jocelyne, 2004, « Mnémème, Antitopic : le post-Rhème : de l'énoncé au texte », in Fernandez-Vest, MM J. – Shirley Carter-Thomas (éds.) *Structure informationnelle et Particules énonciatives. Essai de typologie*, L'Harmattan, Coll. Grammaire et Cognition, 1-2 : 63-100.
- FLEISCHMANN Suzanne, 1982, *The Future in Thought and Language : Diachronic Evidence from Romance*. Cambridge, Cambridge University Press.
- FRADIN Bernard, 2003, *Nouvelles approches en morphologie*, Paris, PUF.
- FRANCOIS Jacques (dir.), 2004, *L'adjectif en français et à travers les langues*, Presses Universitaires de Caen.
- FUCHS Volker – Serge MELEUC, 2003, *Linguistique française : français langue étrangère. La communication en français*. Berne, Peter Lang.
- FUKUI Naoki, 1995, « The principles-and-parameters approach : a comparative syntax of English and Japanese », in Shibatani – Bynon, pp. 326 – 346.
- GABELENTZ, von der Georg 1891, 1901² : *Die Sprachwissenschaft, ihre Aufgaben, Methoden und bisherigen Ergebnisse*, Leipzig.
- GREEN, John N., 1988, « Spanish », in Harris–Vincent, *The Romance Languages*, 79–130.
- GREENBERG Joseph H. 1954 ; 1960 « A quantitative approach to the morphological typology of language », *International Journal of*

- American Linguistics* 26 : 178-194 (reprinted from *Method and Perspective in Anthropology* : Papers in Honor of W. D. Wallis, 1954)
- 1963, 1966² « Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements » in Greenberg (éd.) *Universals of Language*, Cambridge, The MIT Press : 73-113.
- (éd.), 1968, *Universals of Language*. Cambridge, MIT Press.
- (éd.) 1978, avec C. A. Ferguson, E. Moravcsik. *Universals of Human Language*, 4 vols. Stanford, Stanford University Press.
- HAJDÚ Péter – Péter DOMOKOS, 1980, *Uráli nyelvrokonaink* [Nos parents de langues ouraliennes] Budapest, Tankönyvkiadó.
- HARRIS, Martin, 1978, *The Evolution of French Syntax. A comparative approach*, London-New York, Longman.
- HARRIS Martin – Nigel VINCENT, 1988, *The Romance Languages*, London–New York, Routledge.
- HASPELMATH Martin 2001, « The European linguistic area : Standard Average European », in Haspelmath et al. pp. 1492-1510.
- HASPELMATH Martin - Ekkehard KÖNIG – Wulf OESTERREICHER – Wolfgang RAIBLE, 2001, *Language Typology and Language Universals. An International Handbook*, Berlin – New York, Walter de Gruyter.
- HEINE Bernd – Ulrike CLAUDI –Friederike HÜNNEMEYER, 1991, *Grammaticalization. A conceptual framework*, Chicago-London, The University of Chicago Press.
- HEINE Bernd – Tania KUTEVA, 2002, *World Lexicon of Grammaticalization*, Cambridge, Cambridge University Press.
- HOPPER Paul – Elisabeth TRAUGOTT, 2003, *Grammaticalization*, Cambridge, Cambridge University Press.
- HUTTERER Miklós, 1986, *A germán nyelvek* [Les langues germaniques], Budapest, Gondolat Kiadó.
- JAKOBSON Roman, 1958, « Les études typologiques et leur contribution à la linguistique historique comparée », in *Essais de linguistique générale*, éd. Fr. Minuit, 1963, pp. 68-77.
- JONES Michael, 1988, « Sardinian », in Harris–Vincent, *The Romance Languages* : 341–350.
- KATAMBA Francis – John STONHAM, 2006², *Morphology*, New York, Palgrave-Macmillan.
- KEENAN, Edward 1976, « Towards a Universal Definition of Subject » in Li Charles – Thompson Sandra (ed.), *Subjects and topics*, New York, Academic Press, 305-333.

- KENESEI István – VÁGÓ Róbert – FENYVESI Anna, 1987, *Hungarian*, London- New York, Routledge.
- KESZLER Borbála (éd.), 2000, *Magyar grammatika* [Grammaire du hongrois], Budapest, Nemzeti Tankönyvkiadó.
- KLAUSENBURGER Jürgen, 2000, *Grammaticalization. Studies in Latin and Romance morphosyntax*, Amsterdam/Philadelphia, John Benjamins.
- KLINKENBERG Jean-Marie, 1994, *Des langues romanes*, Louvain-la-Neuve, Duculot.
- KURYLOWICZ Jerzy, 1965, *The Evolution of Grammatical Categories*, in *Esquisses Linguistiques*, Munich, Fink, II : 38-45.
- LAMBRECHT Knud, 1994, *Information structure and sentence form*, Cambridge, University Press.
- LAZARD Gilbert, 1999, « Comment comparer l'incomparable ? », *LINX* no. spécial, 15-30.
- 2006, *La quête des invariants interlangues. La linguistique est-elle une science ?* Paris, Champion.
- LEHMANN Christian, 1985/1995, *Thoughts on Grammaticalization*, München, LINCOS Studies in Theoretical Linguistics 01.
- LEHMANN, Winfred P, 1973, « A Structural principle of language and its implications », *Language* 49 : 42-66.
- LI Charles –Sandra THOMPSON (éd.), 1976, *Subjects and topics*, New York, Academic Press.
- LYONS John, 1970 (trad. frse.), *Linguistique générale. Introduction à la linguistique théorique*, Paris, Larousse.
- MALHERBE Michel, 1983, *Les langues de l'humanité*, Paris, Seghers.
- MALLINSON Graham, 1988, « Rumanian », in Harris-Vincent, *The Romance Languages*, 391-419.
- MARCHELLO-NIZIA Christiane 2006, *Grammaticalisation et changement linguistique*, Bruxelles, De Boer-Duculot.
- MARTINET André, 1970, *Éléments de linguistique générale*, Paris, Armand Colin.
- MEILLET Antoine, 1948 (rééd. de 1912), *Linguistique historique et linguistique générale*, Paris, Champion.
- MIRA MATEUS, Maria – et al. 1989, *Gramática da Língua Portuguesa*, Caminho SA, Lisboa 1989²
- PARKINSON Stephen, 1988, « Portuguese », in Harris-Vincent, *The Romance Languages*, 131-269.
- PERROT Jean, 1978 « Fonctions syntaxiques, énonciation, information », *BSLP* 73,1 :85-101.
- PITTAU Massimo, 1972², *Grammatica del sardo-nuorese*, Bologna, Pàtron.

- PLANK Frans, 1998, « The co-variation of phonology with morphology and syntax : A hopeful history », in *Linguistic Typology* 2-2, 1998, p. 195-230.
- 1999, « Split morphology : How agglutination and flexion mix », in *Linguistic Typology* 3-3, pp.279-341.
- PLUNGIAN Vladimir A., 2001, « Agglutination and flecion », in *Language Typology and Language Universals*, HASPELMATH et al., Berlin–New York, De Gruyter, vol. I., pp. 669-678.
- POP Sever 1948, *Grammaire roumaine* Bern, Francke.
- RASTIER François 2001, « Sciences cognitives et Intelligence Artificielle », in *Language Typology and Language Universals*, HASPELMATH et al. Berlin–New York, De Gruyter, vol. I., 2001, pp.75-85.
- RENZI Lorenzo (avec SALVI, Giampaolo), 1989², *Grande grammatica italiana di consultazione*, Il Mulino, Bologna.
- RIEGEL Martin – PELLAT, – RIOUL, R., 1994, *Grammaire méthodique du français*, PUF, Paris.
- ROBINS R. H., 1964, 1973 (trad. frse), *Linguistique générale. Une introduction*, Paris, Armand Colin.
- SAPIR Edward, 1921, *Language, an Introduction to the study of speech*, New York, Harcourt, Brace and World
- 1968. *Linguistique*, Paris, Ed. Minuit.
- SAUSSURE Ferdinand De, 1916, *Cours de linguistique générale*, Paris, Payot, éd. 1990.
- SCHLEGEL August Wilhelm, 1818, *Observations sur la langue et littérature provençales*, Paris.
- SCHLEICHER August, 1850, *Die Sprachen Europas in systematischer Übersicht*, Bonn.
- SCHWEGLER Armin, 1990, *Analiticity and Syntheticity. A Diachronic Perspective with Special Reference to Romance Languages*. Berlin, Mouton de Gruyter.
- SHIBATANI Masayoshi, 1985, « Passives and related constructions: a prototype analysis », *Language*, 61-4:821-848.
- SHIBATANI Masayoshi – Theodora BYNON, 1995, *Approaches to Language Typology*, Oxford, Clarendon Press.
- SÓRÉS Anna 1989, « Esquisse d'une typologie synchronique des langues romanes », *Revue de Linguistique Romane*, 209-210 ; 5-24.
- 1993, « La place des langues romanes dans la typologie syntaxique moderne », *Actes du XXe Congrès International de Linguistique et Philologie Romanes*, tome III, section IV : 359-368.

- 1995, « Rapports génétiques et typologiques dans l'étude synchronique des langues romanes », *Revue Romane*, 30/1 : 41-79.
- 1999 a, « Topique, focus et ordre des mots en hongrois », *La thématization dans les langues*, Actes du Colloque de Caen 1997, éd. Claude Guimier, Peter Lang, Berne, 217-230.
- 1999 b, « Le hongrois, langue SOV ou SVO : problèmes dans l'établissement de l'ordre de base », *LINX Numéro spécial*, Actes du colloque *Typologie des langues, universaux linguistiques*, éd. A. Sörös – C. Marchello-Nizia, 189-204.
- 2001, « Néhány gondolat a kontrasztív nyelvészet és a tipológia kapcsolatáról » [Quelques réflexions sur le rapport entre la linguistique contrastive et la typologie], in *Színes eszmék nem alszának* [Les idées colorées ne dorment pas], Hommage à György Szépe, Pécs, Lingua Franca Csoport, pp. 1146–1153.
- 2003, « A propos du passif en français et en hongrois », in *Frontières et passages*. Actes du colloque franco-hongrois sur la traduction, éd. T. Szende – G. Máté, Coll. Etudes contrastives, Peter Lang, Berne :265-274.
- 2004 a, « La notion "d'ordre de base" dans la typologie des langues et son application à la langue hongroise, *BSLP*, t. XCIX (2004), fasc. I. p. 281-310.
- 2004 b, « La place de l'adjectif épithète dans les langues. Approche typologique » in *L'adjectif en français et à travers les langues*, dir. J. François, Presses Universitaires de Caen, 89–140.
- 2005 a, « La structure informationnelle du hongrois » in *Les langues ouraliennes aujourd'hui : approche linguistique et cognitive* (éd. J. Fernandez-Vest), Paris, H. Champion, Bibliothèque de L'Ecole des Hautes Etudes à la Sorbonne, 517-531.
- 2005 b (avec Christiane Marchello-Nizia) : « Typologie diachronique : une nouvelle hypothèse pour le changement de type 'OV'>'VO' » dans les Actes du 3^{ème} Colloque *Typologie des langues, universaux linguistiques*, Editions du Septentrion, Lille, pp. 261-288.
- 2006, *Le hongrois dans la typologie des langues*, Limoges, Lambert – Lucas.
- 2007 (avant parution), *Typologie et linguistique contrastive. Théories et applications dans la comparaison des langues*, Berne, Peter Lang.
- STEELE Susanne, 1978, « Word order variations : a typological study », in *Universals of Human Language I-IV.*, GREENBERG Joseph H. (ed.), Stanford, Stanford University Press, pp. 587-623
- TEYSSIER, Paul, 1992², *Manuel de langue portugaise*, Klincksieck, Paris.
- TOMPA József, 1968, *Ungarische Grammatik*, The Hague, Mouton.

- TRAUGOTT Elisabeth Closs – Bernd HEINE, 1991, *Approaches to grammaticalization*, Amsterdam/Philadelphie, John Benjamins.
- UNGERER Friedrich – Hans-Jörg SCHMID, 1996, *An Introduction to Cognitive Linguistics*, London-New York, Longman.
- VAZQUEZ CUESTA, Pilar – Maria A MENDES da Luz,, *Gramática da língua portuguesa*, Edições 70, Lisboa 1980.
- VENNEMANN Theo, 1973, «Explanation in syntax »; *Syntax and semantics* 2, ed. John K. Kimball, New York, Seminar Press : 1-50.
- WAGNER, M. L. (1951) *La lingua sarda*, Bern, Francke.
- WAGNER, R. L. – PINCHON, Jacqueline, 1962, *Grammaire du français classique et moderne*, Hachette, Paris.
- WHALEY Lindsay J., 1997, *Introduction to typology. The unity and diversity of language*, Thousand Oaks – London – New Delhi, SAGE Publications.
- WHEELER Max W. – Alan YATES– Nicolau DOLS, 1999, *Catalan. A Comprehensive Grammar*, Routledge, London–New York.