

La morphologie

La morphologie étudie les variations de forme des mots régulièrement associées à des différences de signification. Le concept de morphème est central dans la plupart des approches modernes de la morphologie, mais il ne permet d'expliquer de façon simple qu'une partie des phénomènes que doit prendre en considération une étude de la structure interne des mots, et il a été fortement remis en question à date récente par un certain nombre de théoriciens.

Bien que la morphologie soit depuis toujours une composante essentielle de la linguistique descriptive, son histoire récente n'a pas été marquée par des débats autour de textes théoriques de façon comparable à ce qui a pu être observé en phonologie ou en syntaxe. Les débats théoriques concernant la morphologie se sont souvent résumés à une discussion sur l'incidence que des théories essentiellement phonologiques ou syntaxiques peuvent avoir sur la morphologie.

Haspelmath (2002) *Understanding morphology* (Arnold Publishers, Londres) fournit à la fois une introduction générale très accessible et une excellente présentation des questions théoriques qui sont au centre des préoccupations actuelles des spécialistes de morphologie.

Les notions essentielles de la morphologie

Mot et lexème : Nous ne donnons pas à 'mot' le même sens lorsque nous disons par exemple que le mot *dormez* comprend deux syllabes, ou que *dormez* et *dort* sont deux formes du même mot. On utilise *lexème* pour se référer spécifiquement à cette dernière valeur (ce qui est commun à un ensemble de mots regroupés sous une même entrée lexicale¹), et *mot-forme* pour préciser que c'est à la première valeur qu'on se réfère (le mot en tant que séquence de phonèmes ou de lettres). Un mot-forme est un fragment d'énoncé qui se caractérise à la fois par un haut degré de cohésion interne et par une relative autonomie par rapport aux autres mots avec lesquels il entre en combinaison². On désigne comme *paradigme* l'ensemble des mots-formes considérés comme appartenant à un même lexème. Le *sens lexical* d'un mot est ce qui est sémantiquement commun à l'ensemble des mots-formes qui constituent le paradigme auquel il appartient. Une *famille de mots* est un ensemble de lexèmes présentant des ressemblances partielles relativement régulières à la fois dans leur sens et dans leur forme³.

Mots pleins et mots grammaticaux : Le test le plus généralement invoqué pour opérer cette distinction est que les mots pleins constituent des classes ouvertes et peu stables, alors que les mots grammaticaux constituent des classes fermées et relativement stables, mais l'application de ce critère est parfois problématique, et la question de sa justification théorique reste entière. Une définition particulièrement intéressante consiste à poser que les classes de mots pleins sont celles où vont se ranger les lexèmes qui sémantiquement ont un sens dénotatif indépendant de toute référence à la construction d'un discours (par exemple, un objet est reconnaissable comme *table* ou *chaise*

¹ Ceci n'est malheureusement pas la seule acception avec laquelle on peut rencontrer le terme de lexème : chez certains auteurs, lexème est synonyme de mot plein, et la notion de lexème exclut les mots grammaticaux, alors que selon la définition retenue ici elle les englobe ; chez d'autres, le terme de lexème s'applique à des unités significatives minimales (ou morphèmes lexicaux, par opposition aux morphèmes grammaticaux).

² L'apprentissage de l'écrit développe chez les locuteurs d'une langue l'idée que le découpage des phrases en mots est quelque chose qui va de soi. En fait, le découpage orthographique de la phrase écrite en mots est largement conventionnel et va souvent à l'encontre des décisions que l'on devrait prendre si on découpait les phrases en appliquant de manière cohérente des critères précis.

³ Dans la description d'une langue, le paradigme des formes fléchies d'un lexème doit être délimité de façon parfaitement nette ; il ne serait par contre pas raisonnable d'avoir une exigence comparable pour la délimitation des familles de mots.

quelles que soient les conditions dans lesquelles on en parle), alors que la signification des mots grammaticaux est relative à la construction d'un discours (par exemple, il n'y a pas d'objet qui soit intrinsèquement *celui-ci* ou *le mien*). Cette définition ne résoud pas toutes les difficultés, mais elle permet au moins d'éviter un certain nombre de faux problèmes.

Beaucoup de difficultés auxquelles on se heurte dans la description des langues tiennent au fait que sont constamment à l'œuvre des processus de *grammaticalisation* par lesquels les mots pleins présentant certaines caractéristiques sémantiques tendent à devenir des mots grammaticaux, et les mots grammaticaux tendent à perdre leur autonomie et à devenir des affixes.

Flexion et dérivation : La *morphologie dérivationnelle* étudie la relation entre lexèmes considérés comme appartenant à une même 'famille de mots', et la *morphologie flexionnelle* étudie la relation entre les différentes formes d'un même lexème. Il est facile de reconnaître comme relevant de la dérivation les processus morphologiques qui changent la catégorie syntaxique des mots (*chanter* — > *chanteur*), par contre la question des critères permettant de caractériser comme flexionnels ou dérivationnels les processus morphologiques qui ne changent pas la catégorie syntaxique des mots est une question délicate. Un critère particulièrement utile fait référence à l'interaction entre la structure morphologique du mot et son comportement syntaxique : seules les caractéristiques flexionnelles du mot peuvent intervenir dans la formulation de règles syntaxiques, et il ne doit pas exister de règle syntaxique spécifique à tel ou tel type de mot dérivé. Autrement dit, la dérivation peut modifier les propriétés syntaxiques du mot, mais le comportement syntaxique d'un mot dérivé doit pouvoir se prédire en le rattachant simplement à une catégorie à laquelle appartiennent aussi des mots non dérivés. Par contre, une règle de syntaxe peut comporter des conditions nécessaires sur les caractéristiques flexionnelles des mots auxquels elle s'applique⁴.

Morphèmes, morphes et allomorphes : Dans les cas simples, comme celui de l'exemple suivant, on peut rendre compte de la structure morphologique des mots en découpant les mots-formes en segments qui véhiculent chacun une partie de la signification.

<i>aim-er</i>	<i>aim-ons</i>	—	<i>aim-able</i>	—
<i>trouv-er</i>	<i>trouv-ons</i>	—	<i>trouv-able</i>	<i>in-trouv-able</i>
<i>lav-er</i>	<i>lav-ons</i>	<i>lav-eur</i>	<i>lav-able</i>	—
<i>cass-er</i>	<i>cass-ons</i>	<i>cass-eur</i>	<i>cass-able</i>	<i>in-cass-able</i>

Il est tentant de poser comme principe général que, de même que la phrase est une chaîne de mots, le mot est une chaîne de segments significatifs minimaux, qui se présentent chacun comme une chaîne de phonèmes (ou de lettres), mais il n'y a rien d'exceptionnel à ce que les relations régulières entre mots prennent des formes qui se prêtent mal à une telle analyse, et il est préférable de reconnaître la possibilité d'opérations morphologiques autres que la simple concaténation de segments. Et même dans les cas où la concaténation de segments suffit à rendre compte de la structure des mots, il est utile de distinguer entre le *morphème* (plus petite partie significative d'un mot isolable par segmentation) et le *morphe* (manifestation phonique ou graphique du morphème). Deux segments significatifs minimaux (morphes) sont *allomorphes* d'un même morphème abstrait s'ils expriment la même signification et apparaissent en distribution complémentaire. Par exemple,

⁴ Selon ce critère, l'infinitif du français est une formation flexionnelle, car les propriétés syntaxiques des infinitifs ne se retrouvent dans aucun mot qui ne présente pas la terminaison caractéristique des infinitifs ; par contre, le nom d'action est une formation dérivationnelle, car les noms d'action ont les mêmes propriétés syntaxiques que les noms qui ne sont pas dérivés de verbe. Le temps verbal est une catégorie flexionnelle, car aucun mot dépourvu de variation en temps ne peut se substituer au verbe conjugué. Le pluriel des noms est aussi en français une catégorie flexionnelle, du fait de l'existence de règles d'accord en nombre.

en français, *sai-* (dans *je sais*), *sav-* (dans *nous savons*) et *sau-* (dans *tu sauras*) sont trois allomorphes de la racine du verbe *savoir*.

Les verbes anglais qui diffèrent du nom correspondant par la place de l'accent (*annex* (V) / *annex* (N), *concord* (V) / *concord* (N), *contrast* (V), *contrast* (N), etc.) fournissent un exemple typique de relation morphologique régulière entre mots qui échappe à un traitement en termes de concaténation de morphèmes.

Mots, clitiques et affixes : Une forme (morphème ou séquence de morphèmes) peut être caractérisée comme libre ou liée selon qu'elle peut ou non constituer à elle seule un mot. Une forme libre peut être qualifiée de minimale si toute suppression d'un de ses éléments constitutifs lui enlève son caractère de mot. Plusieurs types de critères peuvent être utilisés pour reconnaître qu'une limite entre deux segments significatifs sépare l'une de l'autre deux formes libres, ou qu'au contraire au moins l'une des deux formes situées de part et d'autre de cette limite est une forme liée :

–un premier type de critère repose sur la notion d'indépendance prosodique des formes libres : une forme qu'il est naturel d'énoncer en isolation, par exemple en répondant à une question, est nécessairement une forme libre ; si une phrase est interrompue par des pauses, les pauses tendent à coïncider avec des limites entre formes libres, et il en va de même pour l'insertion d'incises, etc.

–un deuxième type de critère repose sur le fait que les possibilités d'insertion d'éléments supplémentaires entre un morphème lié et le reste du mot dont il fait partie sont limitées à un petit nombre de morphèmes qui sont eux-mêmes des morphèmes liés : dans *ré-organiser*, le seul élément insérable entre le préfixe *re-* et la base *organiser* est lui aussi un préfixe (*re-dés-organiser*) ;

–un troisième type de critère repose sur l'impossibilité de traiter isolément les morphèmes liés dans des opérations auxquelles se prêtent les mots ou groupes de mots, telles que la coordination ; en français, le fait qu'on ne puisse pas dire **je te et lui parle* ou **tu et je viendrons* montre que les pronoms personnels conjoints, bien qu'écrits traditionnellement comme des mots à part, sont des formes liées.

Mais les morphèmes grammaticaux (ou séquences de morphèmes grammaticaux) reconnaissables selon les critères précédents comme formes liées ne manifestent pas toujours au même degré des caractéristiques d'intégration à un mot qui les engloberait. Certaines formes liées représentent des mots grammaticaux dont la position dans la phrase est déterminée par les règles de la syntaxe, mais qui sont susceptibles de se *cliticiser*, c'est-à-dire de perdre leur autonomie pour former une unité prosodique avec un mot plein auquel ils sont adjacents. C'est par exemple le cas en anglais pour 've forme réduite de *have*, ou 's forme réduite de *is*. Un *clitique* est en principe une forme liée qui se caractérise par un faible degré d'interaction avec son *hôte*⁵, tandis qu'un *affixe* se caractérise par une interaction forte avec la *base* à laquelle il s'attache⁶, mais dans la pratique il s'agit d'une distinction dont l'établissement soulève souvent des problèmes délicats.

Quelques autres définitions à connaître

agglutinant : un système morphologique idéalement agglutinant serait un système morphologique connaissant comme seule opération morphologique la concaténation de morphèmes.

alternance morphophonologique : processus phonologique qui a perdu sa motivation phonétique initiale mais qui reste actif dans la relation entre allomorphes d'un même morphème.

amalgame : processus historique par lequel la limite entre deux morphèmes successifs s'estompe et finit par disparaître.

⁵ Les clitiques se subdivisent en *enclitiques* (qui suivent leur hôte) et *proclitiques* (qui précèdent leur hôte).

⁶ Les principaux types d'affixes sont les *préfixes* (qui précèdent leur base) et les *suffixes* (qui lui succèdent).

- analogie : modification de mots ou création de mots nouveaux sur le modèle de mots déjà existants.
- analytique : une langue analytique est une langue qui tend à exprimer par des constructions syntaxiques ce que d'autres expriment par des opérations morphologiques.
- base : la base d'une forme morphologiquement complexe est l'élément auquel s'applique l'opération morphologique qui donne naissance à cette forme.
- classe flexionnelle : deux lexèmes appartiennent à la même classe flexionnelle s'ils ont le même type de paradigme et si les formes qui constituent leur paradigme résultent des mêmes opérations morphologiques.
- coalescence : processus historique par lequel la combinaison de deux formes libres se fige en un mot unique.
- composé : lexème complexe dans la formation duquel entrent au moins deux autres lexèmes.
- cumul : expression de deux significations morphologiques par un segment significatif minimal (non décomposable comme la concaténation de deux segments significatifs plus petits).
- défectif : un lexème défectif est un lexème qui ne possède pas toutes les formes fléchies que possèdent normalement les lexèmes de la même catégorie.
- désinence : équivalent d'affixe flexionnel, dans une tradition terminologique qui réserve le terme d'affixe à la morphologie dérivationnelle.
- fusionnant : un système morphologique fusionnant fait une large place à des opérations qui ne se ramènent pas de façon simple à des concaténations de morphèmes, et à des phénomènes d'allomorphie qui n'ont pas de traitement phonologique simple.
- isolant : une langue idéalement isolante serait une langue analytique à l'extrême, dans laquelle chaque morphème aurait le statut de mot.
- morphophonologie : l'étude des alternances morphophonologiques.
- morphème vide (ou morphème zéro) : on peut reconnaître dans un mot-forme la présence d'un morphème vide lorsque le statut de ce mot-forme dans le paradigme auquel il appartient se reconnaît à l'absence du morphème qui marque le statut des autres mots-formes du paradigme.
- périphrase : construction syntaxique utilisée en concurrence avec une formation morphologique.
- porte-manteau : un morphe porte-manteau est un segment significatif minimal qui cumule deux significations morphologiques distinctes.
- productif : un modèle morphologique est dit productif s'il peut s'appliquer à de nouvelles bases pour former de nouveaux mots.
- racine : base irréductible (non analysable comme étant elle-même le résultat d'une opération morphologique).
- réanalyse : processus historique par lequel un mot est interprété comme le résultat d'opérations morphologiques qui ne sont pas celles qui ont initialement servi à le former.
- réduplication : opération morphologique qui diffère de l'affixation proprement dite par le fait que ce qu'on ajoute à une base est une copie (totale ou partielle) de cette même base.
- représentation sous-jacente : représentation abstraite de la structure du mot, qui ne correspond pas de manière immédiate à la forme phonique ou graphique, mais que le linguiste postule pour pouvoir formuler avec un maximum de simplicité les règles qui génèrent les différents mots-formes d'un même paradigme, ou les différents lexèmes d'une même famille.
- soustraction : opération morphologique consistant à supprimer une partie d'une base.
- supplétion (ou supplétivisme) : utilisation d'une forme inanalysable à la place du produit régulier d'une opération morphologique.
- synthétique : une langue synthétique est une langue qui tend à exprimer par des opérations morphologiques ce que d'autres expriment par des constructions syntaxiques.
- thème : base à laquelle s'appliquent directement les opérations de la morphologie flexionnelle.

La syntaxe

1. Quelques repères théoriques. La syntaxe est traditionnellement définie comme l'étude des régularités auxquelles se conforme la combinaison des mots en phrases. Après une période (il y a quelques décennies) où la distinction traditionnelle entre syntaxe et morphologie était fortement remise en question, elle est à nouveau considérée comme cruciale.

Au 19^{ème} siècle et pendant la première moitié du vingtième siècle, les travaux théoriques d'envergure sur la syntaxe ont été relativement peu nombreux, l'attention des linguistes soucieux de jeter les bases d'une description scientifique du langage (et pas seulement de poursuivre dans la voie tracée par la grammaire traditionnelle) se portant surtout sur la phonologie et la morphologie. Vers la fin des années 50, les premiers travaux de Chomsky ont radicalement modifié la situation, et depuis, les débats théoriques sur la syntaxe ont pris une importance considérable.

On classe souvent par commodité les théoriciens actuels de la syntaxe en 'formalistes' ou 'généralistes' (Chomsky) et 'fonctionnalistes' (Givón, Dik, Van Valin), mais il s'agit d'une dichotomie trompeuse. En effet, le fait de viser à une description formalisée de manière plus ou moins précise est en principe indépendant des grandes options théoriques sur le statut de la syntaxe dans le fonctionnement du langage. Il est parfaitement possible d'accepter les postulats des 'fonctionnalistes' tout en ayant un souci de formalisation équivalent (sinon même supérieur !) à celui de bien des 'formalistes' (c'est notamment l'ambition de théories 'fonctionnalistes' telles que celle de Dik, ou celle de Van Valin). Le débat fondamental est sur le fait de concevoir la syntaxe comme un système autonome, directement conditionné par des aptitudes inscrites spécifiquement dans le patrimoine génétique humain (c'est le postulat défendu par Chomsky), ou au contraire de penser qu'il n'y a qu'une relation relativement indirecte entre le système syntaxique et les aptitudes cognitives de l'être humain, la syntaxe n'étant qu'un élément parmi d'autres dans un système langagier globalement organisé en fonction des besoins communicatifs de l'être humain et des conditions dans lesquelles se déroule la communication par le langage. Il est d'ailleurs parfaitement possible d'estimer que l'état actuel de nos connaissances (aussi bien en linguistique que dans d'autres domaines) ne permet pas d'avancer dans ce débat autrement que de façon purement spéculative, et donc qu'il est souhaitable de travailler sur la description syntaxique des langues et sur la typologie syntaxique (comme le font d'ailleurs en réalité la plupart des syntacticiens, y compris parmi ceux qui prennent explicitement position dans le débat) d'une façon qui ne présuppose pas de s'être préalablement rangé dans un camp ou dans l'autre.

Sur un plan plus 'technique', il y a une distinction importante entre les chomskyens et la quasi-totalité des autres écoles (y compris celles se revendiquant généralistes à l'égal du chomskysme) sur la question de savoir s'il convient de décrire directement les contraintes combinatoires qui font que seules certaines séquences de mots constituent des phrases bien construites, ou s'il est préférable de les décrire de façon indirecte, par référence à des structures abstraites hypothétiques ayant les caractéristiques suivantes : elles peuvent comporter des éléments 'invisibles', les unités n'y sont pas forcément rangées selon l'ordre 'de surface', et les morphèmes liés peuvent dans ces structures abstraites fonctionner comme entités syntaxiques autonomes.

Pour une présentation détaillée des différents domaines de la syntaxe avec une forte orientation typologique, cf. Denis Creissels, *Cours de syntaxe générale* (sera publié comme livre en 2004 ou 2005, mais sera disponible sur le web, pour les étudiants de Lyon 2, dès janvier 2004)

Les notions essentielles de la syntaxe

1. La phrase comme structure dans le cadre de laquelle s'expriment systématiquement les modalités de l'énonciation. Si on ne veut pas se contenter d'une approche intuitive, la définition de

la phrase doit s'appuyer sur les notions de *contenu propositionnel* et d'*opération énonciative*. Un contenu propositionnel est la représentation d'un état possible du monde (événement, situation) au moyen des lexèmes d'une langue. La notion d'opération énonciative développe les notions sur lesquelles se base le classement traditionnel des phrases en déclaratives (ou assertives), interrogatives et impératives (ou injonctives).

Le propre du langage est de fournir aux locuteurs la possibilité de construire une variété illimitée de contenus propositionnels et d'explicitier les opérations qu'ils effectuent en manipulant ces contenus propositionnels dans l'interaction communicative, et la phrase est le cadre dans lequel se manifestent systématiquement ces possibilités. Les phrases des ex. (1) à (3) ainsi que les groupes de mots entre crochets dans l'ex. (4) renvoient à la même représentation d'un événement *réparer* mettant en jeu *Jean* et *la voiture*, et les variations d'un exemple à l'autre expriment de manière systématique différentes possibilités de manipuler ce contenu propositionnel dans un acte d'énonciation :

–les phrases (1) sont des phrases indépendantes assertives, positive (a) et négative (b) ;

- (1) a. Jean a réparé la voiture
b. Jean n'a pas réparé la voiture

–les phrases (2) sont des phrases indépendantes injonctive (a) et interrogatives (b-d) ;

- (2) a. Que Jean répare la voiture !
b. Qui est-ce qui a réparé la voiture ?
c. Qu'est-ce que Jean a réparé ?
d. Est-ce que Jean a réparé la voiture ?

–les phrases (3) expriment différentes façons de présenter le même contenu informatif en fonction du contexte discursif (ici, il s'agit précisément de *focalisation*) ;

- (3) a. C'est Jean qui a réparé la voiture (... moi, j'en aurais été incapable)
b. C'est la voiture que Jean a réparé (... pas la moto)

–enfin, les groupes de mots entre crochets dans l'ex. (4) sont des constituants phrastiques qui participent à la construction de phrases complexes.

- (4) a. Je crois [que Jean a réparé la voiture]
b. Je ne sais pas [si Jean a réparé la voiture]
c. J'ai demandé à Jean [de réparer la voiture]
d. Jean a promis [de réparer la voiture]
e. Tu peux partir avec la voiture [que Jean a réparée]

La phrase est une combinaison de mots dont la structuration permet l'expression systématique de l'élaboration énonciative d'un contenu propositionnel ; le propre d'une phrase est de participer à un jeu de correspondances régulières avec d'autres phrases qui expriment une élaboration énonciative différente d'un même contenu propositionnel.

2. Les espèces de mots (ou 'parties du discours'). Une répartition des mots en classes est nécessaire pour pouvoir formuler les règles selon lesquelles, dans une langue, certaines combinaisons de mots sont des phrases possibles, et d'autres pas.

Les mots peuvent être classés selon leur *distribution*, c'est-à-dire en observant le fait que seuls certains mots peuvent occuper une position donnée dans une construction donnée :

Cet homme —	Cet homme est —
dort	*dort
*médecin	médecin
*intelligent	intelligent

Les mots peuvent être classés selon les possibilités de faire varier leur structure interne, et certaines au moins des caractéristiques morphologiques du mot ont un lien direct avec son comportement syntaxique : en français, le fait que *dort* et *intelligent* ont une distribution différente est corrélé au fait que *dort* peut varier en personne (*dort / dormons*) ou en temps (*dort / dormait*), alors que *intelligent* n'a pas ces possibilités, tandis qu'inversement, *intelligent* peut varier en genre (*intelligent / intelligente*), mais pas *dort*.

—> La question de la relation entre classes grammaticales de mots et types de significations est à poser une fois le classement grammatical des mots établi selon des critères formels (il n'y a pas une relation nécessaire entre le comportement grammatical et le sens dénotatif des mots pris individuellement, même si au niveau du lexique pris globalement on peut reconnaître des affinités entre certaines classes de mots et certains prototypes sémantiques).

—> Le classement grammatical des mots peut varier de manière imprévisible d'une langue à l'autre : rien ne permet de prévoir qu'en français, *pause* fonctionne exclusivement comme nom, alors que le correspondant anglais de ce mot, *break*, peut fonctionner aussi bien comme nom que comme verbe (*Usually we break at ten = D'habitude nous faisons une pause à dix heures*). Autre exemple : en français, *français* et *allemand* ont exactement le même statut (mêmes possibilités d'emploi à la fois comme noms et comme adjectifs), et rien ne permet de deviner qu'en anglais, *German* peut s'utiliser au singulier comme nom au sens de 'personne de nationalité allemande', alors que la même chose n'est pas possible pour *French* (pour 'personne de nationalité française', on ne peut avoir que *Frenchman/woman*).

3. Constituants syntaxiques. Les locuteurs ont l'intuition de l'existence de 'groupes de mots' qui fonctionnent comme un seul bloc à un certain niveau de la structure de la phrase, et cette intuition peut être précisée par différents tests, notamment en observant des modifications de l'unité phrastique qui impliquent la présence de mots ou groupes de mots en une position différente de leur position canonique :

- Jean a parlé [de son projet de thèse] [au directeur du département]
- C'est [de son projet de thèse] que Jean a parlé — au directeur du département
- C'est [au directeur du département] que Jean a parlé de son projet de thèse —
- *C'est [au directeur] que Jean a parlé de son projet de thèse — du département
- * C'est [de son projet] que Jean a parlé — de thèse au directeur du département
- * C'est [de thèse au directeur] que Jean a parlé de son projet — du département

En dehors des tests reposant sur la possibilité de déplacer en bloc les groupes de mots qui forment des constituants, les manipulations suivantes sont généralement considérées comme susceptibles de fournir des indices de l'existence de constituants syntaxiques :

- possibilité / impossibilité de ne pas répéter certains mots ou groupes de mots (ellipse) : *Si Jean veut [venir avec nous au cinéma], il peut — ;*
- possibilité / impossibilité de substituer à un groupe de mots une proforme : *Jean m'[en] a déjà*

parlé, [de son projet de thèse] ;

–possibilité de coordonner un constituant avec d'autres constituants du même type : *Jean veut [vendre sa vieille voiture] et [acheter une moto].*

4. La structure interne des constituants : tête et dépendants. Parmi les mots qui forment un constituant, la tête est celui qui détermine le statut syntaxique du constituant pris en bloc. Ainsi, le constituant nominal a en principe pour tête un nom, et la phrase a en principe pour tête un verbe. Les autres mots ou groupes de mots qui entrent dans la formation du constituant peuvent être désignés du terme générique de *dépendants* (on dit parfois aussi *modifieur*).

La notion de tête est fondamentalement le reflet syntaxique du fait que, sémantiquement, les dépendants précisent un sens dénotatif qui est donné par la tête : *le nouveau directeur du département* dénote un individu catégorisé comme *directeur*, et *Jean m'a parlé de toi* se réfère à un événement catégorisé comme *parler*.

5. Types majeurs de constituants

Les constituants nominaux ont pour tête un nom : *J'ai vu [la nouvelle voiture de Jean].*

Les constituants phrastiques sont des fragments de phrase ayant eux-mêmes une structure interne de type phrastique (cf. ex. 4 de la section 1)

Les constituants prépositionnels (ou selon les langues, postpositionnels) formés par une préposition et un constituant nominal (ou par un constituant nominal et une postposition) : *Nous nous sommes rencontrés [devant [la nouvelle poste]]*

Les constituants adjectivaux et adverbiaux sont formés par un adjectif ou un adverbe accompagné d'un nombre variable de modifieurs : *Jean est [très fier de sa nouvelle voiture], Jean court [presque aussi rapidement que Paul].*

Le groupe verbal est formé par le verbe et les termes de sa construction autres que le sujet : *Jean [parle [de ses dernières vacances] [à Marie]].*

6. Aspects fondamentaux de la construction de la phrase simple : rôle syntaxique et rôle sémantique des constituants nominaux, et la notion de valence verbale. L'aspect le plus fondamental de l'organisation syntaxique d'une langue est la construction de phrases combinant un verbe avec un certain nombre de constituants nominaux ou prépositionnels qui occupent chacun un rôle syntaxique précis : un terme de la construction du verbe peut être obligatoire ou facultatif, il peut donner lieu ou non à un accord avec le verbe (cf. la notion de sujet), avoir plus ou moins de possibilités de se déplacer, etc.

Il est crucial de ne pas confondre les *rôles syntaxiques* des termes de la construction d'un verbe (notion strictement formelle) avec les *rôles sémantiques* qu'on peut reconnaître à leur référent : les deux couples de constituants nominaux de [*L'enfant*] a cassé [*l'assiette*] et [*L'enfant*] aime [*le chocolat*] ont les mêmes rôles syntaxiques (sujet / COD), mais ils représentent un *agent* et un *patient* dans la première phrase, un *expérient* et un *stimulus* dans la deuxième. La notion de *valence verbale* est cruciale pour comprendre comment s'articulent ces notions : chaque verbe détermine à la fois les rôles sémantiques possibles pour les termes de sa construction et les rôles syntaxiques correspondant à chacun de ces rôles sémantiques.

—> Chaque langue organise les propriétés de valence de ses verbes d'une façon qui lui est propre : En français, *L'enfant a cassé l'assiette* et *L'enfant aime le chocolat* ont la même organisation syntaxique (*sujet - verbe - COD*), alors que les phrases correspondantes de l'espagnol sont syntaxiquement très différentes (*El niño rompió el plato / Al niño le gusta el chocolate*).

Typologie et universaux

L'organisation des langues varie beaucoup plus que ne peut l'imaginer quelqu'un qui n'a aucune expérience du système de langues génétiquement et géographiquement diverses, mais cette variation a ses limites (certains types d'organisation *a priori* concevables ne sont attestés par aucune langue connue). L'objectif de la recherche d'universaux de langage est de préciser ces limites, et la typologie s'efforce de dresser un inventaire aussi synthétique que possible des types d'organisation possibles pour une langue. Typologie et recherche des universaux sont deux aspects complémentaires, et indissociables l'un de l'autre, de l'étude de la diversité des langues.

Au 19^{ème} siècle, la typologie a voulu réduire la diversité des langues à un petit nombre de types (trois dans la version la mieux connue de ce classement : isolant / agglutinant / fusionnant), fondés sur des critères essentiellement morphologiques (pour les définitions, cf. *La morphologie*), avec l'idée que l'ensemble du fonctionnement de la langue devait être corrélé de façon simple aux critères morphologiques pris pour base du classement.

Sapir (*Language*, 1921) a été le premier à montrer que d'une part la typologie morphologique est trop complexe pour se réduire aux trois types classiquement reconnus, et que d'autre part il n'y a pas de relation évidente entre les particularités morphologiques d'une langue et sa façon de construire les phrases. Il a ainsi fondé la typologie moderne, qui se fixe dans un premier temps de recenser point par point les aspects de l'organisation d'une langue pouvant donner lieu à variation, pour poser seulement dans un deuxième temps le problème de l'existence de corrélations.

Greenberg, dans son livre de 1966 *Universals of language*, approfondit cette orientation et jette les bases d'une véritable typologie syntaxique, qui a connu depuis un développement considérable. Comrie (1981) *Language universals and linguistic typology* et Croft (1990) *Typology and universals* sont deux ouvrages fondamentaux pour se faire une idée générale des orientations actuelles et des principaux résultats d'ores et déjà acquis. Pour la typologie phonétique, l'ouvrage fondamental est Ladefoged & Maddieson (1995) *The sounds of the world's languages*.

La variation phonétique. Les sons que l'appareil phonatoire humain peut produire et que l'oreille humaine peut discriminer ne sont pas également attestés dans les langues. Parmi les voyelles, les antérieures non arrondies (i, e) et les postérieures arrondies (u, o) sont beaucoup plus fréquentes que les antérieures arrondies (y, ø) ou les postérieures non arrondies (inexistantes en français). Parmi les consonnes, les langues ont généralement des inventaires de plosives plus importants que leurs inventaires de fricatives, et certaines langues n'ont pas du tout de fricatives. Les plosives p, t, k (ou leurs correspondantes voisées b, d, g) sont attestées dans pratiquement toutes les langues, ce qui est loin d'être le cas pour la plosive q. Certains types de plosives (aspirées, glottalisées) sont propres à certaines langues. Les clics sont particulièrement rares. Parmi les fricatives, s est de loin la mieux attestée dans les langues du monde. Parmi les semi-voyelles, y et w sont très communes, □ (attestée en français) est très rare. Parmi les nasales, m et n sont beaucoup mieux attestées que □ ou ≠.

La variation phonologique. Deux langues qui ont les mêmes sons ne les utilisent pas nécessairement de la même façon : en français, r (apical) et R (postérieur) sont deux variantes régionales ou individuelles du même phonème, alors qu'en portugais ce sont deux phonèmes différents – cf. ['karu] 'cher' / ['kaRu] 'voiture'.

Un aspect important de la variation phonologique est la variation dans les contraintes combinatoires, notamment dans les possibilités de combiner les phonèmes en syllabes. Par exemple, la plupart des langues interdisent des séquences consonantiques comme tk ou pt en début de mot, mais le russe ignore cette interdiction. Le cas extrême de restriction aux séquences de phonèmes est celui de langues où CV est le seul type syllabique possible.

La variation morphologique. Le degré de complexité morphologique des mots varie (analytique vs synthétique), ainsi que les types de significations exprimés au niveau

morphologique : en hongrois, la suffixation de *-hat/het* au verbe équivaut en français à la combinaison de deux mots distincts : *lát-lak* 'je te vois' / *lát-hat-lak* 'je peux te voir'. Il y a aussi des variations dans les manifestations de la structure morphologique des mots complexes : on parle de morphologie agglutinante lorsque chaque élément constitutif du mot correspond à un segment bien identifiable, de morphologie fusionnante pour tout ce qui ne se ramène pas à une simple concaténation de morphèmes. Selon les langues, les éléments grammaticaux du mot peuvent apparaître de façon prédominante comme préfixes (langues bantoues) ou comme suffixes (turc, hongrois).

Quelques aspects importants de la variation syntaxique. Les langues diffèrent tout d'abord dans la façon dont sont mis en forme des contenus propositionnels semblables : disposition linéaire des constituants syntaxiques, marquage de leurs rôles, accord du verbe avec les termes essentiels de sa construction. Par exemple, étant donné un verbe d'action dont le sens implique la participation d'un agent et d'un patient, la façon de ranger le verbe et les termes de la phrase représentant l'agent et le patient varie d'une langue à l'autre : lorsqu'agent et patient sont tous deux définis, le rangement le moins marqué discursivement (parfois le seul possible) peut être *Agent Verbe Patient* (français), *Agent Patient Verbe* (turc), *Verbe Agent Patient* (nahuatl), *Verbe Patient Agent* (malgache).

Kadın kitabı aldı (turc)
femme livre.ACS prendre.PAS.3S
La femme a pris le livre

Kikwa in pilli in nakatl (nahuatl)
3S.3S.manger DÉF enfant DÉF viande
L'enfant mange la viande

Manasa ny lamba ny vehivavy (malgache)
laver DÉF linge DÉF femme
La femme lave le linge

Ensuite, le verbe peut porter des marques renvoyant à certaines caractéristiques de l'agent et/ou du patient, et qui peuvent éventuellement suffire à le représenter.

Arrantzaleak arraina harrapatu du / Harrapatu du (basque)
pêcheur.ERG poisson attraper AUX.3S.3S attraper AUX.3S.3S
Le pêcheur a attrapé un poisson Il/elle l'a attrapé(e)

Arrantzaleek arraina harrapatu dute / Harrapatu dute
pêcheurs.PL.ERG poisson attraper AUX.3P.3S attraper AUX.3P.3S
Les pêcheurs ont attrapé un poisson Ils/elles l'ont attrapé(e)

Arrantzaleak arrainak harrapatu ditu / Harrapatu ditu
pêcheur.ERG poisson.PLattraper AUX.3S.3P attraper AUX.3S.3P
Le pêcheur a attrapé des poissons Il/elle les a attrapé(e)s

Arrantzaleek arrainak harrapatu dituzte / Harrapatu dituzte
pêcheurs.PL.ERG poisson attraper AUX.3P.3P attraper AUX.3P.3P
Les pêcheurs ont attrapé des poissons Ils/elles les ont attrapé(e)s

Enfin, le contraste agent / patient peut se manifester par des marques sur les constituants qui les représentent, et la comparaison de constructions transitives et intransitives montre que ces marques n'ont pas la même distribution d'une langue à l'autre. Par exemple, le russe utilise une marque

morphologique spéciale pour le patient (accusatif), alors qu'en basque, c'est l'agent qui prend une forme différente de celle du mot cité en isolation ou sujet d'un verbe intransitif (ergatif).

Maša rabotaet (russe)
Macha travailler.3S
Macha travaille

Sobaka laet
chien aboyer.3S
Le chien aboie

Maša moet sobaku
Macha laver.3S chien.ACS
Macha lave le chien

Sobaka kusaet Ma%ou
chien mordre.3S Macha.ACS
Le chien mord Macha

Jon eta Mikel etorri dira (basque)
Jon et Mikel venir AUX.3P
Jon et Mikel sont venus

Jonek Mikel ekarri du
Jon.ERG Mikel amener AUX.3S.3S
Jon a amené Mikel

Mikelek Jon ekarri du
Mikel.ERG Jon amener AUX.3S.3S
Mikel a amené Jon

Il y a aussi variation dans le fait d'utiliser ou non la construction typique des verbes d'action impliquant un agent et un patient pour parler d'événements ne mettant pas en jeu les mêmes rôles sémantiques : pour asserter une relation de possession, certaines langues (le basque, le français, ...) ont une construction où le couple possesseur / possédé est traité de la même façon qu'un couple agent / patient, d'autres (le russe, le latin, ...) utilisent une construction différente.

Jonek zakurra du (basque)
Jon.ERG chien avoir.3S.3S
Jon a un chien

U Sa%oi est' sobaka (russe)
près+de Sacha est chien
Sacha a un chien

En ce qui concerne enfin les aspects de la construction des phrases concernant non pas le sens dénotatif, mais l'articulation discursive, on remarque par exemple que les langues diffèrent beaucoup dans la possibilité de marquer la focalisation d'un terme de la phrase par un changement de l'ordre des mots ne s'accompagnant d'aucun autre changement dans la construction de la phrase.

Marco me l'ha detto (italien)
Marco me l'a dit

Me l'ha detto Marco

C'est Marco qui me l'a dit (**Me l'a dit Marco* serait en français une phrase mal construite)

Types de langues et types d'organisation, et la notion de corrélation typologique. Comme cela a été dit dans l'introduction, la typologie moderne a renoncé à définir d'emblée des types de langues, et cherche dans un premier temps à établir les types d'organisation possible dans tel ou tel aspect de leur système pour chercher ensuite des corrélations. Deux exemples de corrélations :

–Les langues dans lesquelles le verbe se place systématiquement en début de phrase utilisent majoritairement des prépositions, celles où le verbe se place systématiquement en fin de phrase utilisent majoritairement des postpositions.

–Les langues à verbe final ont en règle générale une morphologie essentiellement suffixale.

Universaux et tendances, la notion de type marqué, et l'explication des tendances typologiques. A côté d'universaux absolus (par exemple : aucune langue n'utilise la position du verbe en fin de phrase comme marque d'interrogation), beaucoup d'universaux sont seulement des tendances. Par exemple, il existe quelques exceptions à la généralisation selon laquelle aucune langue ne construit ses phrases avec PAV ou PVA comme ordre de base ; on dit que par rapport aux autres types de rangement, PAV et PVA sont des types marqués. En morphologie, la préfixation est typologiquement plus marquée que la suffixation, car si beaucoup de langues ont exclusivement des suffixes, très peu ont exclusivement des préfixes, etc. En phonologie, l'utilisation du trait de voisement est banale pour les consonnes obstruantes, exceptionnelle pour les non-obstruantes.

En ce qui concerne l'explication des tendances typologiques, il faut distinguer plusieurs niveaux. Le seul assez facilement accessible et relativement peu controversé est celui de l'explication historique : l'organisation des langues telles que nous les connaissons est le résultat d'une très longue histoire, or en phonologie aussi bien qu'en grammaire, on observe que la variété des changements possibles est relativement limitée, et que parmi les types de changements possibles, certains se produisent beaucoup plus fréquemment que d'autres. Reste à expliquer ce qui limite la variété des types de changements possibles. En phonétique-phonologie, on peut se baser sur l'idée que tout système phonologique doit représenter un compromis acceptable (mais par nature instable) entre coût articulatoire et rendement distinctif des sons. Pour les systèmes grammaticaux, c'est dans la façon dont fonctionne le processus de production-interprétation des énoncés qu'il faut chercher la motivation des changements.

Typologie et diachronie. On peut se demander dans quelle mesure des affinités typologiques entre langues sont ou non révélatrices d'une parenté génétique. Il convient à ce sujet d'être très prudent, car non seulement les contacts favorisent les évolutions convergentes (linguistique aréale) mais aussi le caractère particulièrement fréquent de certains types de changement peut aboutir à des convergences à première vue étonnantes entre langues très éloignées (le français s'est constitué à date récente un système de pronoms clitiques qui ressemble de plus en plus aux systèmes de préfixes pronominaux du nahuatl ou des langues bantoues).

Typologie et description des langues. La typologie des langues et l'étude des universaux ne peuvent progresser que sur la base d'une information de plus en plus large et de plus en plus précise sur les langues particulières, mais inversement, une information sur les types d'organisation possibles des langues et les problèmes d'analyse qu'ils peuvent poser est indispensable aux linguistes décrivant des langues particulières. Il y a de ce point de vue un problème très sérieux avec l'accélération actuelle de la disparition de langues peu ou pas du tout documentées.