

LES THÉORIES SUR LE VOCABULAIRE

Eléments pour une synthèse

Jacqueline BASTUJI

1. — L'embarras des pédagogues.

1.1. Rénover l'enseignement du français, c'est aussi refuser ou à tout le moins questionner le cloisonnement qui sévit dans l'institution scolaire : cloisonnement synchronique qui découpe la discipline en pratiques disjointes, grammaire, vocabulaire, orthographe, lecture expliquée, et cloisonnement diachronique qui réserve l'enseignement de la langue au premier cycle, et celui de la « littérature » au second.

Dans leur effort pour associer étude méthodique de la langue et maîtrise des textes, les enseignants rencontrent deux domaines où s'observent électivement les phénomènes de contact et d'interaction entre éléments et règles de la *langue* d'une part, et d'autre part ce que Saussure nommait *parole*, et que de nos jours on appelle plus volontiers, et selon la doctrine de référence, *performance*, *texte* ou *discours*. Ces deux domaines sont les modes et temps verbaux, dont les règles d'emploi s'appliquent au-delà des limites de la phrase comme unité canonique de la syntaxe, et le vocabulaire, que les spécialistes nomment *lexique* en vertu d'une distinction sur laquelle nous reviendrons. Or, si les recherches contemporaines — BENVENISTE et WEINRICH — sur les rapports entre temps, personnes et organisations discursives commencent à être assez largement vulgarisées, et notamment par les bons soins de *PRATIQUES*, le lexique demeure le parent pauvre de la pédagogie, et un champ d'incertitude, voire d'exclusion, dans la théorie linguistique contemporaine.

1.2. Qu'il y ait des contraintes dans le maniement des mots de la langue, nul n'en doute, et moins que quiconque les enseignants, qui passent un temps considérable à rectifier les erreurs de l'oral ou à porter en rouge, dans les marges des copies, des appréciations telles que « incorrect », « imprécis », « familier », « impropre », « faux sens », « faute de construction ». Pédagogie strictement négative, et fonctionnant au coup par coup. Peut-on mieux faire, et définir un nombre limité de principes et de règles qui s'appliqueraient à de vastes ensembles de vocabulaire, et aux rapports systématiques entre mot et phrase, mot et contexte ? C'est bien là tout le problème pédagogique.

Si la tradition scolaire fait du vocabulaire un apprentissage distinct, les grammaires récentes qui au nom de la linguistique moderne prétendent y intégrer le lexique lui réservent en fait la portion congrue. Comment, dans ces conditions, s'étonner que les enseignants se réfugient dans l'empirisme bricoleur ? Au hasard des textes, on expliquera, ou on fera expliquer, le sens de *parcimonieux*, de *pittoresque*, de *laser*, de *démocratie* ou d'*orgueil*, mais sans toujours préciser s'il s'agit d'un emploi dans le texte ou d'une valeur dans la langue. Parfois, on fera chercher des synonymes dont on appréciera collectivement les nuances, puis on demandera aux élèves de les employer dans des phrases de leur cru, mais sans disposer d'une théorie de la synonymie et de la polysémie. On fera rectifier * *je me rappellerai longtemps de cette belle journée*, mais sans traiter de façon systématique les schèmes de construction du verbe à l'intérieur de la phrase. Le pourrait-on d'ailleurs ? L'enseignement du vocabulaire se fait par la méthode tâtonnante des essais et erreurs. Plus encore, il se fait par imprégnation : imprégnation par les lectures, l'audio-visuel, les propos tenus en classe, les conversations dans le milieu familial et dans celui des copains. Les enfants des milieux sociaux défavorisés y sont nécessairement perdants, et nous voici renvoyés au vaste problème politique de l'école qui reproduit l'inégalité et la justifie par un système d'évaluation chiffrée des performances langagières. Si l'on refuse le pessimisme radical de ceux qui considèrent qu'aucun changement n'est possible à l'école avant la révolution prolétarienne, et si l'on considère le travail enseignant comme l'un des champs de la lutte entre l'ancien et le nouveau, peut-on trouver, à ces incertitudes, d'autres remèdes que la bonne volonté et l'inspiration du moment ?

1. 3. A première vue, la linguistique semblerait cautionner le bricolage hasardeux des pédagogues, puisque les propriétés qu'elle reconnaît au lexique sont précisément l'hétérogénéité et l'irrégularité : hétérogénéité des niveaux de langue et strates de vocabulaire, infinie polysémie des mots, irrégularité des relations formelles entre les mots simples et leurs dérivés. La seule discipline pertinente serait-elle la *lexicographie*, c'est-à-dire la confection de dictionnaires classant commodément et selon l'arbitraire alphabétique la diversité des formes et des emplois ? En fait, même la lexicographie s'inspire d'une théorie sur les mots. En présentant ici les méthodes et les problèmes caractéristiques des principales écoles qui depuis près d'un siècle travaillent sur le lexique, nous tenterons de montrer qu'elles aboutissent à un grand nombre de controverses sur des problèmes non résolus, mais aussi à un ensemble de propositions dès maintenant exploitables.

2. — Approches lexicologiques et relations entre mot et discours, vocabulaire et société.

2. 1. La lexicologie à dominante externe.

2. 1. 1. La tradition linguistique la plus ancienne s'en tient à l'approche philologique qui caractérise encore les concours de recrutement du Second Degré. Les mots sont analysés à partir des textes, en tant qu'ils livrent des informations sur la culture et l'expérience historique du peuple ou du groupe social qui les utilisent. La systématisation des données ainsi recueillies a donné naissance à la *lexicologie* comme discipline sociologique fondée sur l'analyse des ensembles de mots « dont la structure et les relations seront

expliquées par les faits sociaux » (G. MATORE, La méthode en lexicologie, 1953). Vaste programme, impliquant le développement massif d'une théorie des formations sociales qui est encore largement programmatique, et débordant de loin les objectifs propres à la linguistique.

Appliquée à l'enseignement, une telle lexicologie cautionne les pratiques d'étude de vocabulaire au fil des textes, mais encourage aussi — et cette attitude-là est plus neuve — la relativité critique vis-à-vis de la norme et des niveaux de langue. En tout cas, on demeure dans ce que Saussure appelait la *linguistique externe*, consacrée précisément aux rapports entre langue et société. SAUSSURE n'en conteste pas l'utilité, mais lui reproche d'« accumuler détail sur détail sans se sentir serrée dans l'étau d'un système », et lui oppose la *linguistique interne* comme seule discipline véritablement scientifique : « la langue est un système qui ne connaît que son ordre propre ». Le problème est donc de savoir si le lexique échappe au système de la langue, exclusion que pour sa part SAUSSURE rejetait catégoriquement, ou s'il est possible de l'y intégrer en construisant des procédures d'analyse qui relèvent, soit de la linguistique interne, soit de l'interaction entre linguistique interne et linguistique externe.

2. 1. 2. Justement, le *Cours de Linguistique Générale* proposait deux hypothèses d'une remarquable fécondité. 1° La langue est un système synchronique dont les unités se déterminent par le seul jeu de leurs rapports internes. 2° Pour chaque unité, ces rapports sont doubles : les uns s'établissent horizontalement sur l'*axe syntagmatique* des enchaînements, et les autres verticalement sur l'*axe paradigmatique* des associations, c'est-à-dire des substitutions virtuelles. Tout mot, ou tout emploi particulier d'un mot, se définit à la fois par sa *distribution*, qui est la classe de ses environnements de gauche et de droite, et par les *relations d'opposition pertinente ou d'équivalence* qu'il entretient avec les mots ou séquences qui peuvent lui être substitués dans un environnement identique.

2. 2. La lexicologie structurale et l'analyse de discours.

La lexicologie structurale a travaillé principalement sur des « *champs* », qualifiés de « sémantiques » ou de « lexicaux ». Elle étudie comment les unités s'y déterminent réciproquement, mais à l'intérieur d'un domaine dont l'unité est « notionnelle », et donc extérieure à la linguistique : couleurs, termes de parenté, « vocabulaire de l'aviation » (GUILBERT 1965), « vocabulaire politique et social en France de 1869 à 1872 » (DUBOIS 1962). La problématique ancienne des rapports entre faits linguistiques et sociaux n'est pas abandonnée, mais se fonde désormais sur des procédures rigoureuses de description synchronique, qui en outre font avancer la connaissance du lexique dans ses règles d'organisation formelle et sémantique : matrices de dérivation préfixale ou suffixale, relation d'opposition ou d'équivalence entre des mots comme *avion/aéroplane*, ou *ouvrier/travailleur/prolétaire*, processus néologiques.

De tels travaux partent toujours d'un corpus (un ou plusieurs textes considérés comme homogènes), et les mots y sont analysés à partir de leur contexte : *contexte étroit* de la distribution dans le syntagme ou dans la phrase, mais aussi, éventuellement, *contexte large* du paragraphe ou de l'ensemble discursif, voire *contexte situationnel* regroupant les conditions d'émission et de réception. Or, le recensement des liaisons entre les mots et leurs contextes établit que les locuteurs et les groupes sociaux se distinguent moins par leur vocabulaire que par le réseau des phrases dans lesquelles ils l'insèrent.

Le corpus des journaux parisiens sur la guerre d'Algérie (D. MALDIDIER) ne permet pas d'opposer un « vocabulaire de droite » à un « vocabulaire de gauche ». Le « vocabulaire politique d'Eluard » (M.-R. GUYARD) se fonde sur un réseau de métaphores intéressant les cycles naturels et la végétation : ARBRE, FEUILLE, FLEUR, BOURGEON, GRAINE, FORET, SOLEIL, NUIT...

Les champs sémantiques reposeraient ainsi sur l'illusion présaussurienne d'une relation stable entre signifiant et signifié, et leur pertinence se limiterait à quelques domaines où la terminologie est strictement codée : mythes, institutions, vocabulaires techniques et scientifiques. Aussi le travail lexicologique, et notamment en France, tend-il actuellement à s'épanouir, ou à se diluer, en analyse de discours. Rappelons que par opposition au *texte*, conçu comme message clos défini par sa seule organisation interne, le *discours* englobe la problématique des rapports entre énoncé et énonciation ; « le discours, c'est l'énoncé considéré du point de vue du mécanisme discursif qui le conditionne » (L. GUESPIN, *Langages* 23, p. 10). Par un mouvement de spirale, nous voici ramenés à la linguistique externe, mais à un niveau très supérieur intégrant quelques acquis fondamentaux : données de l'énonciation, variation lexicale selon l'environnement, construction des idéologies comme discours et comme interdiscours.

2.3. La sémantique structurale et la sémiotique textuelle.

C'est sur la langue qu'entendait travailler la sémantique structurale, afin d'y définir et d'y théoriser les relations sémantiques entre des mots substituables dans un même environnement. Sur le modèle de l'analyse du phonème en traits distinctifs, le sens du mot est décomposé en unités différentielles, appelées *sèmes*, dont les uns seraient spécifiques, par exemple [\pm Liquide], [\pm Juridique], [\pm Comestible], [\pm A dossier], [\pm A poils courts], tandis que d'autres, dits *classèmes*, par exemple [\pm Animé], [\pm Concret], seraient assez généraux pour s'appliquer à l'ensemble du lexique. Ces traits s'articulent sur des axes d'opposition, dits *axes* ou *catégories sémantiques*, où ils prennent la valeur positive, négative ou neutre. Sur l'axe sémantique du Sexe, *fille* a le trait [+ Femelle], *garçon* le trait [+ Mâle] (l'attribution de la valeur positive au sexe mâle ou femelle est bien entendu conventionnelle, ou idéologique !, l'important étant de formaliser la polarité), tandis qu'*enfant* neutralise l'opposition du Sexe. La relation sémantique fondamentale est l'*opposition pertinente*, qui associe une ressemblance (les sèmes communs) à une différence (les sèmes différentiels) : *astronome/astronaute*, *amener/apporter*. L'*antonymie* relie deux termes dont tous les sèmes sont communs, sauf un dont les valeurs sont inversées : *chaud/froid*, *grand/petit*. L'*hyponymie* définit une implication non réciproque entre un terme précis, par exemple *tulipe* ou *violette*, et un terme plus général comme *fleur* : toute tulipe est une fleur, mais toute fleur n'est pas une tulipe. Les sèmes du terme général regroupent ce qu'Aristote appelait *le genre commun*, et ses hyponymes leur ajoutent des sèmes spécifiques. La *synonymie* est un cas limite entre des mots dont tous les sèmes seraient communs, et se réduit ordinairement à une *paronymie* avec un ou deux sèmes de différenciation : *dangereux/périlleux*, *chaussures/godasses*.

Telle est l'*analyse sémique* (POTTIER, GREIMAS), dont l'intérêt pédagogique semble indéniable : dépistage des ressemblances et des différences, explicitation des relations régulières entre les mots, entraînement à la pensée

logique et à la mathématique des ensembles. Son domaine de pertinence est cependant limité aux traits les plus simples et les plus universellement répandus. Comment analyser exhaustivement, sans tomber dans la paraphrase lexicographique, des mots comme *macadam* ou *magie*? Et comment réduire à quelques traits stables l'infinie complexité de l'expérience? Peut-on valablement décider si *l'orage* est concret ou abstrait? L'analyse sémique est guettée par le mécanisme et la réduction logico-mathématique, et même en intégrant le concept de *connotation*, elle achoppe sur la métaphore, la polysémie, l'ambiguïté. D'où la tentation de l'appliquer aux textes plutôt qu'à la langue, et en faisant l'économie d'un niveau d'organisation essentiel qui est celui de la phrase. Cette sémantique-là ignore la syntaxe, et se convertit — le cas est particulièrement net chez GREIMAS — en sémiotique textuelle où règne le concept d'*isotopie*, « ensemble redondant de catégories sémantiques qui rend possible la lecture uniforme du récit ». De ce fait, la plupart des analyses sémiques proposées aux enseignants, comme celle d'un très beau poème de MICHAUX dans *Pratiques* 9, ne sont qu'une forme supérieure de l'explication de textes.

3. — Le lexique comme niveau spécifique d'organisation de la langue.

Que l'on prenne les analyses distributionnelles, illustrées notamment en France par les travaux du Laboratoire de M. GROSS, ou que l'on aborde les hypothèses avancées par la grammaire générative dans l'effervescence de ses révisions théoriques et de ses querelles entre écoles rivales, la nouveauté fondamentale tient à la liaison désormais établie entre lexique et grammaire. Les unités lexicales se définissent par les schémas de phrases dans lesquelles elles peuvent s'insérer, et le *lexique* devient un système de virtualités définies par des règles combinatoires, par opposition au *vocabulaire*, sous-ensemble du lexique fondé sur un corpus et se prêtant à inventaire. Les vocabulaires sont multiples et clos, le lexique est un mais regroupe plusieurs centaines de milliers d'unités avec leurs règles de formation et d'emploi. Cette opposition lexique/vocabulaire correspond donc à une opposition plus générale entre langue et parole, compétence et performance.

3.1. Les analyses distributionnelles

Les mots s'analysent par leur combinatoire syntagmatique, c'est-à-dire par les classes de termes et syntagmes avec lesquels ils sont en construction. Ces classes sont définies par une propriété syntaxique, Groupe nominal, Nom, Déterminant... et si besoin est par les traits sémantiques qui contraignent la sélection. Soit les phrases :

Jean / le mur / le vin / le bois travaille.

Jean travaille dans le jardin / à l'atelier / toute la journée / comme un fou.

Jean travaille le bois / la terre / son violon.

Cette histoire me travaille.

Jean travaille à sa thèse / à ma perte / à me nuire.

★ *Le mur / le vin travaille la terre / la fermentation.*

★ *Le mur / le vin travaille à sa perte / à sa dilatation.*

Le recensement systématique des emplois possibles et interdits de *travailler* établit : 1° que par son environnement de droite il admet une triple construction, intransitive, transitive, Groupe prépositionnel à + GN. 2° que cette triple construction est contrainte par des restrictions sélectives sur les groupes nominaux associés. Par exemple, la construction transitive ou par GPrép est incompatible avec un GN de gauche (le groupe sujet) qui n'aurait pas le trait Humain ou à tout le moins Animé (afin d'admettre *Le castor travaille le bois / à construire un barrage*. Le signifiant *travailler* peut ainsi être scindé en plusieurs homonymes dont chacun correspond à un environnement spécifique.

Ce travail de classification ne se fonde pas sur un corpus clos, mais sur l'activité du locuteur — linguiste, enseignant, élève et groupe d'élèves — qui explore toutes les potentialités de l'unité. Il s'accompagne de manipulations qui permettent de repérer les vraies structures par delà la diversité des apparences. Ainsi, *Jean travaille toute la journée* n'est pas une construction transitive en tant que le GN supporte une opération de déplacement qui serait impossible avec *Jean travaille la terre*. De même, à *l'atelier* n'a pas les mêmes propriétés grammaticales que *à sa thèse*. Un tel travail associe le vocabulaire, la syntaxe et l'orthographe, et incite les élèves à construire par eux-mêmes des tableaux de classement qui systématisent les définitions et exemples des dictionnaires, qu'ils pourront consulter après coup. Cette exploration syntagmatique sera complétée par un jeu de substitutions paradigmatiques dégageant les oppositions pertinentes et les paronymies (*cette histoire me travaille / tracasse / préoccupe*). L'analyse sémique est par là soumise à une contrainte distributionnelle qui filtre les emplois et guide leur classement méthodique.

3.2. La grammaire générative et transformationnelle

3.2.1. La « théorie standard »

La grammaire générative et transformationnelle s'est construite comme dépassement critique de la linguistique distributionnelle dont cependant elle entend conserver les acquis, à savoir le classement formel des schémas de phrase. Ces schémas ne sont plus conçus comme le terme de l'analyse, mais comme des *structures de surface* que la grammaire dérive par des règles transformationnelles à partir des structures syntagmatiques de base, dites *structures profondes*.

Le centre de la grammaire étant la syntaxe, le lexique y apparaît d'abord comme le résidu contingent et non analysable des « irrégularités de la langue ». Aveu d'impuissance auquel remédiera le maître ouvrage de 1965, *Aspects de la Théorie syntaxique*, qui tente de formaliser la correspondance entre combinaisons syntaxiques permises et propriétés spécifiques des « items lexicaux ». Chaque item est un symbole complexe regroupant une combinaison de traits, les uns grammaticaux (*traits catégoriels* comme Nom ou Verbe, *traits de sous-catégorisation* comme Transitif), les autres spécifiant les contraintes de sélection, comme [\pm Animé] ou [\pm Dénombrable].

Tout en soulevant maint problème que nous n'évoquerons pas ici, cette théorie du « composant lexical » permet d'intégrer les travaux antérieurs sur l'analyse sémique et les formules distributionnelles. Les sèmes de POTTIER ou de GREIMAS, qui se veulent sémantiques, concordent avec les traits « syntaxiques » proposés par les distributionnalistes et les générativistes. Simplement, ces traits sont intégrés dans des schémas explicites de phrase, et

une distinction est faite entre les *traits inhérents*, qui n'intéressent que le sens de l'unité (par exemple la relation hyponymique entre *tulipe* et *fleur*), et les *traits contextuels* qui règlent la sélection réciproque des items. Enfin ces traits, ou tout au moins les plus généraux d'entre eux, s'ordonnent selon des *règles de redondance* : lorsqu'un mot porte le trait [+ Humain], il a nécessairement les traits Animé, Concret et Dénombrable. *Le lexique est un ensemble composite, mais structuré.*

3. 2. 2. La Sémantique générative

Dans la théorie standard, la sémantique n'est qu'une composante interprétative s'appliquant sur les structures profondes produites par le schéma syntaxique abstrait et les insertions lexicales ; en particulier, toutes les concordances entre les classes grammaticales et les propriétés même sémantiques des mots sont déjà définies dans la structure profonde. L'originalité de la Sémantique générative est de contester cette priorité du composant syntaxique, et de proposer des structures de base sémantiques, ou plus exactement syntactico-sémantiques, sur lesquelles s'appliqueraient des transformations variées et assez lâches aboutissant aux structures syntaxiques de surface. Dans cette optique, l'unité lexicale n'est plus une donnée empirique de départ, un élément intangible venant s'insérer dans un arbre syntagmatique tout prêt à l'accueillir ; l'unité lexicale est un produit de surface où s'est opérée la synthèse entre des unités sémantiques sous-jacentes hiérarchisées en arbre. Ainsi, un mot apparemment simple comme *tuer* dériverait de quatre phrases hiérarchisées (α CAUSE (γ DEVENIR ((NON) VIVANT))) , où α et γ sont les « arguments » nominaux.

Ne pouvant ici exposer les raisons de nos désaccords avec la Sémantique Générative, nous préférons nous limiter à deux aspects positifs de ses propositions sur le lexique. Tout d'abord, celui-ci cesse d'être le parent pauvre de la grammaire puisqu'il relève des mêmes structures complexes et formalisables que la syntaxe. Ensuite, il est clairement montré que les restrictions sélectives de mot sont largement sémantiques, et sont donc sensibles aux facteurs culturels et discursifs. Des phrases comme *Mon poisson rouge me prend pour un imbécile* ou *Ma poêle à frire me prend pour un imbécile* sont parfaitement grammaticales pour des locuteurs animistes, ou dans un discours de fiction.

L'intérêt pédagogique du modèle apparaît clairement dans l'article de M. CHAROLLES sur les verbes de communication (*Pratiques* 9). Ceux-ci sont regroupés selon leur structure syntactico-sémantique, avec le verbe-prédicat qui se construit avec trois arguments : l'Agentif (celui qui parle), l'Objectif (ce qu'on dit), et le *Datif* (le destinataire). En outre, le choix du verbe de communication articule une distinction entre le *contenu posé*, et le contenu présupposé, à savoir un jugement implicite de l'énonciateur sur le message. Avec *dévoiler* ou *apprendre* on présuppose que la communication est véridique, avec *prétendre* qu'elle est erronée, avec *avouer* ou *confesser* qu'elle peut être dangereuse pour le locuteur...

3. 2. 3. La thèse lexicaliste

Répondant à la polémique déclenchée par la Sémantique Générative, CHOMSKY a modifié sa position sur le lexique dans un long article de 1968, *Remarques sur les Nominalisations*, où l'essentiel porte sur l'importance relative du composant lexical et du composant transformationnel. Ou bien l'on considère les mots comme des unités irréductibles dont on note la matrice de

traits distinctifs réglant leur insertion dans les structures profondes ; dans ce cas, le lexique n'est pas touché par les transformations, qui ne s'appliquent qu'aux structures abstraites (passivation, relativisation, etc.). Ou bien l'on considère que l'ensemble des mots — hypothèse forte avancée par la Sémantique Générative —, ou du moins les mots dérivés et composés sont produits par transformation d'une structure de base.

Appliquée à la formation des mots composés et dérivés, cette hypothèse transformationniste a été largement diffusée par des manuels récents qui, même à l'école primaire, demandent aux élèves de reconstituer les « phrases de base ». Le principe d'analyse est le suivant. Soit la phrase, apparemment simple :

La démolition de l'immeuble commencera demain.

Au lieu d'envisager un rapport morphologique, interne au lexique, entre le verbe *démolir* et son dérivé nominal *démolition*, on dérive le Nom *démolition* d'une transformation de nominalisation s'appliquant à l'ensemble de la « phrase de base » *on démolit l'immeuble*, phrase qui se trouve « enchâssée » dans une « phrase matrice » selon le schéma de structure profonde :

P → GN + GV
 [[on démolit l'immeuble] commencera demain]]

Et comme il est fâcheux de présenter aux élèves de pseudo-phrases françaises où des fragments de la langue sont monstrueusement accouplés à des transcriptions métalinguistiques, les manuels remplacent cette structure d'enchâssement par une juxtaposition de phrases du type :

On démolit l'immeuble : cela commencera demain.

où le pronom *cela* a pour fonction de représenter, sous forme nominale, la phrase « de base » enchâssée dans la phrase matrice.

Cette procédure est intéressante en ce qu'elle permet :

1°) de relier systématiquement les mots, les syntagmes et les phrases ;
 2°) de repérer les contraintes syntaxiques et les effets sémantiques de l'opération. Les contraintes syntaxiques intéressent les relations entre le verbe nominalisé et les groupes nominaux qui lui sont associés :

les grossistes achètent d'importantes quantités de fruits → *les achats par les grossistes d'importantes quantités de fruits.*

et cette structure peut être comparée au rapport transformationnel entre phrase active et phrase passive correspondante. Les effets sémantiques sont liés aux effacements de marques et de constituants : certains effacements sont facultatifs comme celui du GN Agent (*les achats d'importantes quantités de fruits*), les autres sont obligatoires, ainsi pour les marques de temps, de mode et de personne sur les Noms. Nominaliser une phrase, ce n'est pas seulement faire plus court ; c'est aussi supprimer des informations qu'on peut avoir intérêt à ne pas donner. Avec un GN, il n'est pas nécessaire de préciser qui est responsable de la démolition de l'immeuble, ni s'il s'agit d'un fait ou d'un simple projet. Ainsi s'explique-t-on que dans un exercice qui se veut strictement grammatical, les élèves peuvent avoir du mal à « retrouver » la fameuse phrase de base.

CHOMSKY entend montrer que ces nominalisations, et par extension toutes les formations de mots dérivés et composés, ne satisfont pas à la régularité des transformations syntaxiques pour trois raisons : productivité réduite du processus, irrégularité de la relation sémantique entre le mot simple et ses dérivés, imprévisibilité partielle des réalisations morphologiques.

Abattre, mais *abattage* et *abattement* ; *chanter* / *chant*, *aboyer* / *aboiement*, *articuler* / *articulation*. Quant à l'irrégularité sémantique, précisons que tous ceux qui font l'action de *passer* ne sont pas des *passants* ou des *passeurs* !

La dérivation lexicale est ainsi marquée par des irrégularités, ou plutôt par des « *sous-régularités* », qui la rendent tout à fait spécifique, et donc interne au lexique. Le lexique a son organisation propre, qui ne saurait être confondue avec la syntaxe : telle est la position dite *lexicaliste*. En même temps, les constituants Phrase et Groupe Nominal ont des propriétés syntaxiques suffisamment voisines pour assurer le passage d'une structure à une autre. Notons enfin que le remplacement du concept d'irrégularité par celui de sous-régularité, en même temps qu'il assure les bases d'une description scientifique du lexique, devrait contribuer à la promotion épistémologique de domaines — fondamentaux — où le nécessaire se trouve associé à l'aléatoire.

4. — Synthèse des acquis scientifiques sur le lexique.

Pour rapide et austère qu'elle soit, cette présentation des affrontements linguistiques sur le lexique nous a semblé une initiation indispensable à des enseignants soucieux de réfléchir sur leurs pratiques, et donc de maîtriser l'ensemble des problèmes théoriques qui les sous-tendent. En outre, et sans vouloir verser dans l'éclectisme, le travail déjà accompli aboutit à un nombre assez remarquable de concordances.

4.1. *La compétence lexicale.*

Alors que les structures syntaxiques ou phonologiques sont réductibles à un ensemble clos d'éléments et de règles, le lexique se caractérise d'abord par son ampleur : plusieurs centaines de milliers d'unités (peut-être 500.000 pour le français), et qui ne peuvent être exhaustivement dénombrées en tant que le lexique est un système synchronique soumis à d'incessants remaniements diachroniques. La fabrication des dictionnaires est une tâche indéfiniment recommencée, et un secteur fort productif de l'activité commerciale. Cette masse énorme constitue le *stock lexical*, c'est-à-dire le « trésor » des mots et locutions qui peuvent être reconnues et comprises comme du français, auquel peut s'ajouter, dans la perspective générative, la totalité des morphèmes et des règles de dérivation et de composition lexicales. Le lexique est un *système de virtualités* rendant compte à la fois des mots attestés et des mots possibles, même si, bien entendu, il importe au pédagogue d'établir une frontière entre les emplois communément reçus — les seuls que les élèves ont le droit d'utiliser au moins dans les épreuves « académiques » —, et les emplois incongrus ou néologiques.

De ce système de virtualités chacun ne maîtrise qu'un sous-ensemble qui constitue son ou ses vocabulaires. Le lexique est divisible en strates diversifiées selon la fréquence et les domaines d'expérience. Le *vocabulaire fondamental* se réduit à un petit ensemble de mille à trois mille mots qui sont les plus fréquents et couvrent les besoins minimaux de la vie courante. Le *vocabulaire commun*, qui regroupe de vingt à trente mille mots, correspond à la compétence d'un locuteur raisonnablement cultivé ; c'est pourquoi le *Dictionnaire du Français Contemporain* s'est limité à 25.000 termes, alors que le *Petit Robert*, déjà plus encyclopédique, en propose le double. Au-delà du vocabulaire commun se déploie le fourmillement des *vocabulaires spé-*

cialisés, où s'articule la diversité des pratiques sociales et des situations de communication : langue littéraire, vocabulaires scientifiques, techniques et professionnels, régionalismes, emplois populaires, obscènes, argotiques, etc...

Il en résulte que la *compétence lexicale* est spécifique — alors que les structures syntaxiques sont, pour l'essentiel, acquises vers cinq ans, l'apprentissage du lexique est un travail qui se fait et se défait tout au long de la vie —, et qu'il existe un important décalage entre la *compétence passive* (les mots que l'on comprend) et la *compétence active* (les mots que l'on sait utiliser effectivement). Les enseignants doivent donc, 1° admettre que la maîtrise du lexique, y compris la leur, est toujours partielle et mouvante ; 2° donner aux élèves les moyens de découvrir et de vérifier par eux-mêmes les formes et les emplois des mots (bon usage des dictionnaires, maîtrise des règles combinatoires, curiosité active et critique appliquée aux énoncés) ; 3° distinguer les mots fréquents, sur lesquels il faut porter l'effort principal, et les vocabulaires spécialisés qui s'acquerront partiellement au fil des textes rencontrés ou des centres d'intérêt, scolaires ou extra-scolaires.

Il faut enfin savoir que si les structures syntaxiques sont des schémas opératoires abstraits, les mots sont des formes où s'investissent l'expérience affective et intellectuelle du sujet, et la culture des groupes sociaux et de la communauté linguistique auquel il appartient. Les enfants, et d'ailleurs les adultes, apprennent le sens des mots en les devinant par le contexte, et par le souvenir qu'ils gardent d'emplois antérieurs. Au besoin, ils s'accrochent à des faux sens dans lesquels ils ont logé tout un travail d'imagination : l'écrivain COLETTE raconte ainsi que pendant tout un an, elle fut « curé sur un mur » parce qu'il lui plaisait de croire que *presbytère* était le nom donné à un petit escargot. On ne peut travailler sur les mots sans les raccorder à des contextes discursifs et à des situations de communication, qu'on choisira aussi variés et aussi « motivants » que possible pour l'enfant ou l'adolescent. On ne peut travailler sur les mots sans rencontrer les règles de décodage — la lecture — et d'encodage — la créativité de l'enfant.

4.2. Propriétés des unités lexicales.

4.2.1. Ces règles d'encodage et de décodage relèvent d'abord de la langue, et peuvent être enseignées comme telles. L'unité lexicale est ordinairement un mot, mais peut en regrouper plusieurs selon un double critère sémantique et syntaxique. Critère sémantique fondé sur l'unité de représentation, et mis en évidence par commutation de la « locution » ou de la « *lexie complexe* » avec un mot simple : *immeuble / grand ensemble, fer à souder / tournevis / marteau, salle de bains / cabinet de toilette / cuisine*. Critère syntaxique fondé sur l'inséparabilité totale ou partielle des éléments : * *un fer neuf à repasser, * une salle étroite à manger, * une salle à manger la glace*. Le foisonnement de la *lexie complexe*, surtout nominale, est caractéristique de nos sociétés contemporaines.

4.2.2. Cette unité lexicale, mot ou *lexie*, se définit par un triplet de propriétés, phonétiques, grammaticales et sémantiques, à savoir :

- la forme sonore du mot et sa transcription orthographique ;
- la classe grammaticale et les règles de construction — transitif ou non, type de groupe prépositionnel associé, type de complétive ;
- les traits de sélection (sujet animé ou non, etc...) et les traits sémantiques inhérents dégagés par le jeu des oppositions paradigmatiques.

Une telle définition présente l'intérêt d'ajouter la dimension syntaxique à la définition saussurienne, et déjà stoïcienne, du signe comme association d'un signifiant et d'un signifié. Les propriétés phonologiques sont incluses dans le signifiant ; quant aux propriétés syntaxiques, qui sont essentielles en tant que le mot se définit comme combinatoire, elles sont l'objet d'un débat théorique qui — nous l'avons vu — est loin d'être tranché et concerne les rapports entre syntaxe et sémantique, langue et discours.

4. 2. 3. Homonymie et polysémie.

Ce repérage des propriétés syntaxiques et sémantiques s'opère par un travail systématique sur les combinaisons syntagmatiques et les substitutions paradigmatiques, travail de type saussurien qui est parfaitement accessible aux élèves et leur permet de découvrir et de contrôler l'essentiel de ce qu'il faut savoir sur le lexique. Ce faisant, on peut regrouper des mots qui ont des distributions analogues, et s'organisent en champs sémantiques structurés : verbes de communication, vocabulaire de l'habitation, etc... Inversement, une même forme sonore peut être répartie entre plusieurs homonymes qui se distinguent par leur combinatoire phrastique et dérivationnelle et par leur signification. Il y a trois verbes *raffiner* : un V transitif qui se construit avec un objet [+ Concret] [— Animé] (*raffiner du pétrole, du sucre*) ; un autre V transitif qui se construit avec un objet [— Concret] (*raffiner son langage, son costume*) ; et un troisième qui se construit avec le groupe prépositionnel *sur* + GN (*raffiner sur sa toilette*). Les dérivés nominaux sont *raffinage* et *raffinerie* pour le premier verbe, et *raffinement* pour les deux autres.

Le découpage en homonymes présente un double intérêt pédagogique : il incite les élèves à l'analyse de l'identique qui se révèle être un multiple, ce qui est un excellent entraînement à la lecture et à l'écoute des messages, et il les oblige à définir les règles de la combinatoire de phrase, règles qui restreignent fortement le jeu paradigmatique de la parasynonymie, domaine où les dictionnaires, tant unilingues que bilingues, offrent le plus souvent des listes inutilisables parce que non distributionnelles. En même temps, *ce découpage opératoire en homonymes ne peut venir à bout du processus infini de la polysémie*. Sauf dans le langage scientifique, qui veille à la précision et à la rigueur monosémique de sa terminologie, le sens d'un mot varie toujours en fonction de contexte phrastique et discursif. Les élèves prennent ainsi conscience que le lexique se définit comme une combinatoire syntaxique et sémantique partiellement ouverte, que l'activité métaphorique s'y déploie nécessairement, et que toute lecture ou toute production d'énoncés est à la fois une obéissance aux règles de la langue et une création à partir de ces règles : création qui bien sûr ouvre un large champ à l'activité discursive, littéraire et poétique.

5. — Synthèse des problèmes et des propositions pédagogiques.

5. 1. Problèmes généraux.

Concevoir un apprentissage méthodique du lexique, c'est d'abord admettre qu'un travail réflexif et systématiquement manipulateur sur la langue — travail qui bien entendu concerne aussi bien la syntaxe que le vocabulaire — améliore la compétence des élèves, c'est-à-dire leur pouvoir réel d'écrire et de parler. Or cette hypothèse, qui bénéficie d'un consensus séculaire dans la tradition scolaire occidentale, est actuellement contestée par un

nombre non négligeable de psychologues. Accordons-lui toutefois le bénéfice provisoire du doute, bénéfice bien nécessaire au confort intellectuel des enseignants de tous niveaux, et recommandons à ceux-ci de ne surtout pas confondre le savoir théorique — savoir explicite et métalinguistique — qu'ils possèdent eux-mêmes sur la langue et qui leur permet de conduire leur classe, et la compétence implicite qui pour les élèves demeure l'essentiel.

Ce gros lièvre une fois soulevé puis recouché, il faut encore observer que le travail sur le lexique, domaine de l'irrégularité et de la « sous-régularité », pose des problèmes spécifiques qu'ignore la syntaxe. Prenons le classique exemple de l'exercice « à trous » où l'on remplace le mot manquant par des points de suspension. Ou bien l'élève connaît déjà le mot approprié et l'exercice est inutile, sauf à titre de contrôle des connaissances. Ou bien il ne le connaît pas, et l'exercice est pernicieux en ce qu'il l'incite à commettre une erreur et à la fixer dans sa mémoire : comment, dans le cas d'une nominalisation, saurait-il d'avance si le bon suffixe est *-age*, *-ment* ou *-tion/ation* ? De telles activités n'ont de sens que si elles sont collectives et entraînent le cas échéant le recours au dictionnaire. Plus généralement, comment éviter ces deux écueils opposés et complémentaires que sont la pédagogie du mécanisme — on donne d'avance le modèle de la réponse et l'élève perd son temps à faire de la copie ou de la production orale stéréotypée — et la pédagogie de la faute, tristement illustrée par le séculaire exercice de la dictée ?

Toute pédagogie du vocabulaire doit être fondée sur l'activité du sujet parlant, activité fondée sur des règles que les enfants ne peuvent véritablement maîtriser que s'ils sont amenés à les découvrir et à les expérimenter par eux-mêmes. Telle est la bonne méthode scientifique applicable à toutes les sciences d'observation, c'est-à-dire à un donné empirique multiforme dont on travaille à découvrir et formuler les principes sous-jacents d'organisation. Cette méthode scientifique est parfaitement assimilable par des enfants et des adolescents, qui sont naturellement aptes à observer, à classer, à critiquer, et qui aiment le faire quand il s'agit d'une activité de groupe où l'on met en commun les idées, les questions et les erreurs à rectifier. Classe-laboratoire où s'éprouve la diversité des usages linguistiques, et où le maître n'a pas nécessairement le dernier mot, puisqu'aussi bien toute compétence lexicale est ouverte et lacunaire. Classe-laboratoire où s'ébauche une attitude neuve vis-à-vis du savoir, comme processus inachevé et cependant guidé par une méthode, et vis-à-vis du pouvoir comme échange diversifié entre les membres d'une communauté. Une telle pédagogie est profondément démocratique et comme telle, s'oppose au modèle régnant de l'école comme transmission autoritaire d'un savoir entièrement constitué et programmé, et comme sélection des « aptitudes » par seule conformité à une norme explicite et implicite. Une telle pédagogie est sans doute dès maintenant applicable, mais seulement, hélas !, par une « avant-garde » d'enseignants. Ceux qui acceptent d'« animer » la classe sans vouloir imposer la totalité d'un programme préparé à l'avance, mais sans non plus démissionner devant une « créativité » que les élèves pourraient sournoisement exercer comme désordre, doux chahut ou bavardage. Ceux qui ont assez de connaissances, notamment linguistiques, pour ordonner la recherche, distinguer l'essentiel de l'accessoire, et ne se faire les esclaves, ni de la dernière théorie à la mode, ni du savoir prédigéré que transmettent les manuels. Ceux enfin dont la situation professionnelle est suffisamment assurée pour affronter les critiques de l'administration ou l'inquiétude des parents.

5.2. Diversité des objectifs.

Dans la pratique, et compte tenu de l'immensité et de l'irrégularité du lexique, on se limitera à des questions ponctuelles, mais exemplaires tant par la méthode que par l'étendue du champ lexical ainsi couvert. On ne négligera donc :

— ni le travail en langue portant sur des unités dont on analyse les propriétés syntagmatiques et paradigmatiques à partir de phrases rencontrées ou produites par les élèves ;

— ni l'indispensable liaison entre orthographe et vocabulaire, liaison qui s'opère électivement au niveau de l'unité lexicale définie tant par son signifiant que par ses propriétés combinatoires ;

— ni le travail à la fois syntaxique, lexical et discursif qui étudie les rapports entre les phrases de base et les dérivations lexicales qui en résultent avec leurs déformations systématiques. On renoncera donc à une approche strictement transformationnelle, qui néglige la spécificité de la dérivation lexicale, mais sans retomber pour cela dans le bric-à-brac descriptif des « familles de mots » où domine l'obsession étymologique ;

— ni l'initiation au maniement des dictionnaires, activité fondamentale s'il en est, puisque le dictionnaire synthétise l'ensemble des informations sur le lexique, qu'il est une production socialement contingente et diversifiée selon les publics et selon les objectifs (dictionnaires de langue et dictionnaires encyclopédiques), et qu'enfin chaque locuteur doit savoir, toute sa vie durant, recourir au dictionnaire pour des informations orthographiques, phonétiques, grammaticales ou sémantiques ;

— ni le rapport entre les mots et les choses, rapport dit de *référence* que Saussure voulait exclure de la linguistique, et qui cependant intervient puissamment dans l'apprentissage des mots existants et dans la création des mots nouveaux. En contestant B. RUSSELL qui affirmait que « personne ne peut comprendre le mot *fromage* s'il n'a pas d'abord une expérience non linguistique du fromage », JAKOBSON rappelle que personne ne peut non plus comprendre le sens du mot *fromage* s'il n'a pas une connaissance suffisante du code lexical français. En fait, le double savoir est nécessaire, et le professeur de français, qui est aussi un professeur de culture et de communication, doit aussi veiller à la correspondance entre les mots sur lesquels on travaille et les fragments de l'expérience qu'ils dénotent ;

— ni le travail sur les textes, que ceux-ci soient lus, entendus ou produits. C'est là que s'opère l'articulation entre langue et discours, et que s'organisent des réseaux sémantiques originaux et cependant respectueux des règles de la langue, selon une créativité discursive qui est analysable et intéresse la sémantique, la stylistique, et les échanges et affrontements entre groupes sociaux.

BIBLIOGRAPHIE

1. Manuels scolaires

- M. DAUMAS et R. LAGANE *Comment apprendre le vocabulaire ? Niveau 2*, 1975, Paris, Larousse.
- J. BASTUJII *Comment apprendre le vocabulaire ? Niveau 3*, 1975, Paris, Larousse.
- M. OBADIA, R. DASCOTTE, M. GLATIGNY, L. COLLIGNON *Le lexique*, 1976, Paris, Hachette, « Les chemins de l'expression », classe de 3^e (avec livre du maître).

2. Articles pédagogiques

- S. DELESALLE *Problèmes du lexique*. Revue LE FRANÇAIS AUJOURD'HUI, n° 27, octobre 1974, Paris.
Publié également dans *Enseignement du français et linguistique : problèmes pratiques et théoriques, Premier cycle du Second Degré*. Revues RECHERCHES PÉDAGOGIQUES, n° 63, 1974, Publications de l'I.N.R.D.P., Paris.
- S. DELESALLE *La recherche du « mot juste », son sens et son rôle*. Revue LANGUE FRANÇAISE, n° 26, mai 1975, Paris, Larousse.
- M. CHAROLLES *Exercices sur les verbes de communication*. Revue PRATIQUES, n° 9, mars 1976.
- H. COTTEZ *Enseignement et faits lexicaux*. Revue LANGUE FRANÇAISE, n° 33, février 1977, Paris, Larousse.

3. Ouvrages et articles de référence

A. TEXTES DE BASE

- S. DELESALLE et M. N. GARY-PRIEUR *Le lexique, entre la lexicologie et l'hypothèse lexicaliste*. Revue LANGUE FRANÇAISE, n° 30, mai 1976, Paris, Larousse.
On consultera l'ensemble des articles de ce numéro qui porte pour titre *Lexique et grammaire*.
- H. MITTERRAND *Les mots français*, coll. Que sais-je ?, 1972, Paris, P.U.F..
- A. REY *Le lexique : images et modèles. Du dictionnaire à la lexicologie*, 1977, Paris, Armand Colin.

B. TEXTES COMPLÉMENTAIRES

- J.-P. BOONS, A. GUILLET, Ch. LECLERE *La structure des phrases simples en français — constructions intransitives*, 1976, Genève - Paris, Droz.
- N. CHOMSKY *Aspects de la théorie syntaxique*, 1971, Paris, Le Seuil (ch. 2 et 4).
N. CHOMSKY *Remarques sur la nominalisation*, in « Questions de sémantique », 1975, Paris, Seuil.

- D. CORBIN *Peut-on faire l'hypothèse d'une dérivation en morphologie ?*, in « Grammaire transformationnelle : Syntaxe et lexique », Publications Université Lille III, S.P. 18, 59650 Villeneuve d'Ascq.
- J. DUBOIS *Le vocabulaire politique et social de 1869 à 1872*, 1962, Paris, Larousse.
- J. DUBOIS et Cl. DUBOIS *Introduction à la lexicographie : les dictionnaires*, 1971, Paris, Larousse.
- M. GALMICHE *Sémantique générative*, 1975, Paris, Larousse (ch. 4).
- M.N. GARY-PRIEUR *Déboiser et déboutonner : remarques sur la construction du sens des verbes dérivés par DE*, in « Grammaire transformationnelle : Syntaxe et lexique », op. cit.
- L. GUILBERT *La néologie lexicale*, Langages 36, décembre 1974.
La créativité lexicale, 1975, Paris, Larousse.
- M.R. GUYARD *Le vocabulaire politique de Paul Eluard*, 1974, Paris, Klincksieck.
- P. GUIRAUD *Structures étymologiques du lexique français*, 1967, Paris, Larousse.
- R. JACKENDORFF *Régularités morphologiques et sémantiques*, in « Théorie générative étendue », 1977, Paris, Hermann.
- J. LYONS, *Linguistique générale*, 1970, Paris, Larousse (ch. 10).
- G. MATORE *La méthode en lexicologie*, 1953, Paris, Didier.
- A. REY *La lexicologie, lectures*, 1970, Paris, Klincksieck.
- F. de SAUSSURE *Cours de Linguistique Générale*, 1973, Paris, Payot.