

Analyse et synthèse en grammaire comparée

Monsieur le recteur,
Monsieur le doyen de la Faculté des lettres,
Mes chers collègues,
Mesdames, Mesdemoiselles, Messieurs,

A l'occasion de cette installation officielle, je suis heureux d'exprimer ma reconnaissance aux autorités universitaires et cantonales, qui ont bien voulu me confier l'enseignement de la grammaire comparée des langues indo-européennes. L'honneur qui m'échoit de professer dans cette maison s'associe dans mon esprit au privilège d'y avoir étudié. Il y a, en effet, tout juste vingt ans que je m'inscrivais comme étudiant à l'Université de Neuchâtel. Le programme des cours était alors moins épais, mais la grammaire comparée y figurait déjà. Mes camarades d'études se souviennent comme moi de cette fascination du vendredi matin. M. Georges Redard nous exposait magistralement les fondements de la linguistique indo-européenne et de la linguistique générale. Dans le cours *ex cathedra*, si original et riche de vues personnelles, l'inépuisable enseignement de Ferdinand de Saussure prenait une dimension nouvelle à la lumière des recherches les plus récentes. Dans le cadre plus intime du séminaire, le maître nous initiait tour à tour, avec une ardeur et une sûreté de jugement admirables, aux singularités du vocabulaire latin, à la complexité de la grammaire sanskrite et à l'interprétation linguistique des textes. Ces leçons comptent parmi les grands moments de ce temps d'études. Avec le recul, j'en saisis encore mieux l'importance et la relation avec nos cours de philologie classique et de philologie médiévale, également très chers à ma mémoire. Aussi, je rends hommage à M. Redard du service incomparable qu'il a rendu à des générations d'étudiants.

La grammaire comparée a pour objet la reconnaissance d'un système de rapports entre les langues d'une famille. Dans la pratique, ces rapports s'observent à différents niveaux:

phonologique, lexical et surtout *grammatical*. Ainsi se justifie la dénomination traditionnelle de *grammaire* — et non de linguistique — *comparée*. Mais, dans sa forme reçue, l'expression *grammaire comparée* n'enferme pas une définition explicite de la discipline. Comme la dénomination de la linguistique appliquée, par exemple, laisse précisément dans le vague le domaine d'application, la désignation de la grammaire comparée ne dit rien sur la nature des termes mêmes de la comparaison. La terminologie germanique n'est guère plus éclairante. A quoi s'applique la *vergleichende Grammatik* ou « grammaire comparante » des Allemands ? Que compare au juste le linguiste ? En fait, non pas une grammaire avec un objet extérieur à la langue, mais des grammaires entre elles en vue de la définition d'une parenté linguistique et au moyen de procédures spécifiques. Là est l'essentiel : il n'y aurait pas de grammaire comparée sans une méthode comparative rigoureuse. En effet, l'approche et l'étude des problèmes s'effectuent par une suite de démarches systématiques. On procède d'abord au choix et à la délimitation des éléments à comparer. En tant que sujette aux coïncidences fortuites, la ressemblance des traits généraux de structure n'entre pas en considération, du moins dans un premier temps de la recherche. Le comparatiste s'attache avant tout aux faits singuliers, comme les alternances de formes au sein de la flexion et les procédés particuliers de la formation des mots. Les correspondances les plus probantes ne se constatent pas au plan de l'énoncé, mais dans les termes de la phrase et jusque dans les dernières subdivisions du signe. C'est pourquoi les opérations d'analyse jouent un rôle primordial en grammaire comparée. Cette méthode analytique s'emploie dans l'étude de toute famille de langues, mais avec un succès particulier dans le domaine des langues indo-européennes. Ce groupe occupe une place privilégiée sous le rapport des travaux comparatifs. Cela tient à la fois aux caractères intrinsèques des idiomes, ni trop semblables, ni trop différents les uns des autres, et à l'avancement des recherches. L'attention du comparatiste se porte donc le plus souvent sur les langues indo-européennes et de préférence sur les états anciens de ces langues, car plus l'observateur remonte le cours du temps, plus les concordances interdialectales apparaissent nettes et la parenté étroite. Ainsi, par la force des choses, l'indo-européaniste se trouve privé de l'outil principal du descripteur d'une langue vivante : l'enquête auprès des sujets parlants. Les matériaux bruts de la grammaire comparée sont des textes anciens, plus ou moins nombreux selon les traditions : abondants en sanskrit ou en grec, par exemple, rares en vieux perse ou en gaulois. De langue à langue la nature des témoignages écrits

varie au gré des cultures et suivant les époques. Il y a loin de l'*Illiade* grecque au *Mahābhārata* indien, en dépit de leur commune appartenance au genre épique. Mais la dissemblance des œuvres comme objets littéraires globaux ne condamne pas la comparaison linguistique. Simplement, celle-ci ne porte pas sur des textes entiers, mais sur des unités plus petites: phrases, syntagmes, mots, morphèmes. La délimitation des phrases n'est pas toujours chose facile. Alors que dans la langue parlée l'unité phrastique est signalée formellement par la pause et l'intonation, et qu'elle l'est dans la langue écrite d'aujourd'hui par les signes de ponctuation, elle ne fait l'objet d'aucun indice matériel dans les documents anciens. Tout le problème revient donc à la découverte des articulations de l'énoncé. Or, les instruments de cette découverte sont l'*analyse* et la *synthèse*. On décompose le discours en une suite de phrases par la démarche analytique. Mais, dans la pratique, la compréhension d'un texte ancien suppose la connaissance du lexique et de l'expression des rapports syntaxiques. L'identification de la phrase procède donc non seulement d'une analyse textuelle, mais aussi d'une construction, c'est-à-dire d'une opération de synthèse à partir des signes et des règles de leur combinaison. Parmi les constituants de l'énoncé, une place particulière revient à des termes points de repère, différents selon les langues. Les particules, notamment, remplissent très souvent une fonction démarcative outre diverses fonctions sémantiques (valeurs explicative, adversative, copulative, etc.). Le sanskrit, par exemple, signale par *īti* une phrase au discours direct. Ce mot articulatoire, à la manière des deux points (:) de la ponctuation moderne, marque un changement de registre dans le discours ou le récit. En règle générale, il se place en fin de citation et la proposition introductive précède ou suit. Exemple: RV 8, 100,3: « tel ou tel dit: Indra n'existe pas ». Dans le texte, la structure syntaxique présente l'ordre inverse: Indra n'existe pas + *īti* + tel ou tel dit (*néndro astīti (= asti īti) néma u tva āha*). Comme particule conclusive, *īti* intéresse donc la délimitation des phrases. C'est le cas aussi d'un groupe de monosyllabes à emploi restreint dans l'énoncé: Véd. *īd, u, gha (ghā, ha), nú, sma, svid, hí* apparaissent presque toujours en deuxième position. Ces éléments sont sémantiquement pauvres — la plupart renforcent seulement le mot précédent —, mais ils ont un rôle syntaxique en tant qu'indicateurs des unités phrastiques. Quelques-uns n'ont pas d'accent propre. Or, Wackernagel, dans un article célèbre, a montré que la tendance existait non seulement en sanskrit, mais déjà en indo-européen commun, de ranger les formes enclitiques à la seconde place de la phrase¹. Le compariste dispose donc de jalons pour l'analyse

d'un texte en unités assertives. Pourtant, les critères sémantiques et formels utilisables ne suffisent pas à la solution de tous les cas. En matière de segmentation, des incertitudes subsistent et les éditeurs d'auteurs anciens hésitent bien souvent entre des ponctuations également défendables. Un exemple de cette difficulté se rencontre, dans la tradition latine, chez Caton, *De agricultura* 5, 3. On y lit à propos du fermier, dans une sorte de cahier des charges et des interdictions: «Qu'il ne fasse pas de sacrifice... sans ordre du maître; qu'il ne prête à personne» (*rem diuinam... ne faciat iniussu domini credat nemini*). Le sens et la syntaxe du morceau autorisent parfaitement une coupe après *faciat*. Cela donnerait alors: «Qu'il ne fasse pas de sacrifice...; qu'il ne prête à personne sans ordre du maître.» Comme l'écriture ne renseigne pas sur les intonations de phrase et comme l'ordre des mots est libre en latin, nul indice matériel ne détermine le rattachement du syntagme *iniussu domini* à la séquence précédente ou suivante. C'est pourquoi les éditions ne concordent pas: le texte de la collection Teubner a un point après *faciat*, le passage correspondant chez Budé un point virgule après *domini*.

Si l'analyse du texte en phrases précède logiquement l'identification des mots, dans la pratique les unités du niveau inférieur se présentent souvent de façon plus immédiate. Par exemple, les graveurs d'inscriptions antiques recourent parfois — et c'est une aubaine pour le déchiffreur — à un signe spécial pour la séparation des mots. Ainsi, le séparateur apparaît régulièrement dans les monuments vieux perses en cunéiforme du VI^e au IV^e siècle av. J.-C. Il s'agit d'un trait oblique ou d'un coin, inscrit avec soin et constance dans l'ensemble de nos documents. Sinon la réalisation matérielle, du moins le procédé se retrouve identique dans la tradition occidentale. Des inscriptions alphabétiques latines marquent nettement les frontières de mots: la fibule de Préneste, des environs de 600 av. J.-C., le fait à l'aide de deux points superposés; la stèle du Forum, du début du V^e siècle, à l'aide de trois points superposés. Donc, les mots sont là, sous la forme de données brutes antérieures à l'analyse. Une pareille commodité rend alors possible une série d'observations précieuses. On relève, en particulier, les structures phoniques admises à l'initiale et à la finale de ces unités libres et on en tire des lois, applicables à l'intérieur de la langue en question. Dans le cas présent du latin, des particularités négatives se dégagent clairement des faits. Certaines combinaisons ne se rencontrent jamais au début ou à la fin d'un mot. Le groupe *sm*, par exemple, n'est admis qu'à l'intérieur. Sur des renseignements de cet ordre, le linguiste fonde ses interprétations ultérieures. Soit un texte en écri-

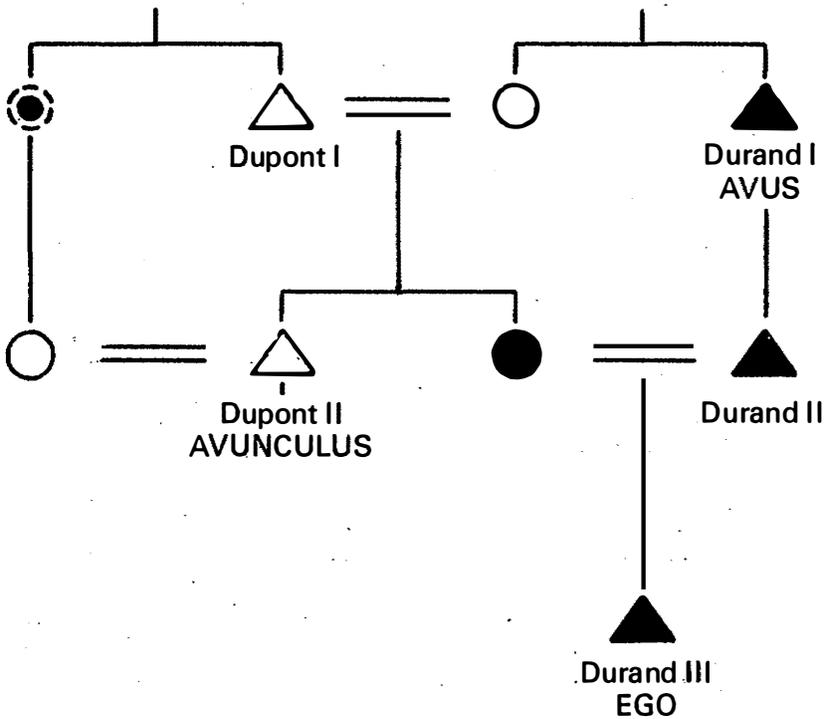
ture continue, la fameuse inscription romaine du Vase dit de Duenos, à peu près contemporaine de la stèle du Forum. Les exégètes y découpent la séquence *duenosmedfeced* en trois mots: *duenos med feced* « bonus me fecit », « un homme bon m'a fait » (c'est en quelque sorte le vase qui parle). Si des considérations de sens, ainsi que la comparaison avec le latin classique, recommandent suffisamment cette segmentation, le critère phonique y contribue aussi. En effet, en vertu du principe de l'exclusion du groupe *sm* des positions initiale et finale, les séquences **dueno/smed* et **duenosm/ed* n'entrent pas en ligne de compte. D'ailleurs, la récurrence, au fil du texte, de segments identiques confirme encore le bien-fondé de l'interprétation. En effet, la forme *med* apparaît non seulement dans *duenosmedfeced*, mais aussi dans les groupes *qoimedmitat* et *dzenoinemedmaaostatod*. Le fait probant est la découverte d'une constante, le monosyllabe *med*, dans des environnements différents. Egalement riche d'enseignements, mais non praticable sur le texte du Vase de Duenos, l'épreuve inverse mettrait en relation le complexe *duenosmedfeced* avec un énoncé comme **duenoshoccefeced* « un homme bon a fait ceci ». La substitution de **hocce* à *med* dans le même contexte met en lumière les articulations de l'énoncé. Mais l'opération de commutation ne donne pas nécessairement de la phrase une analyse en mots. Ainsi, le descendant direct de v. lat. *feced*, le classique *fēcit* « il a fait », coïncide partiellement au plan formel avec *iēcit* « il a jeté ». Or, les éléments interchangeables, *f-* et *i-*, n'ont pas le statut de mots, ni même d'unités significatives. On parle, alors, d'unités distinctives. En revanche, de la comparaison entre *fēcit* « il a fait » et *fēcerunt* « ils ont fait » se dégagent des formes liées porteuses de signification: *-erunt*, par exemple, exprime la troisième personne du pluriel du parfait et *fēc-* véhicule la notion de « faire », mais aussi la valeur du parfait par opposition à *fac-*. Comparer *fēcit* « il a fait » et *facit* « il fait ». L'analyse, on le voit, ne s'arrête pas aux frontières du mot. L'élément *fēc-* du latin, non admis en emploi libre, n'en représente pas moins une réalité pour le linguiste et pour le sujet parlant. Car le sujet parlant pratique aussi une forme d'analyse des unités linguistiques. Preuve en est l'analogie. L'apparition chez l'enfant d'un terme **faisez* dans le paradigme français du verbe « faire » suppose la décomposition de *faisons* en *fais-* + *-ons*. Il y a ensuite identification du *-ons* de *faisons* avec le *-ons* de *pouvons*, par exemple. Enfin, le rapport de *pouvons* à *pouvez* entraîne la paire symétrique *faisons*/**faisez*. La création d'une forme analogique constitue donc un indice de l'analyse du signe linguistique par l'usager. Le Romain avait conscience de l'unité morphologique de *fēc-*, comme le locuteur français reconnaît la

réalité du radical de *faire*. Mais, par l'identification de *fēc-*, le sujet parlant atteint les dernières limites de la décomposition de la chaîne en unités significatives. Le linguiste va plus loin. Comme le physicien poursuit les constituants ultimes de la matière, le grammairien comparatiste démonte les structures compliquées de la racine. Cette opération demande alors la mise en œuvre d'un instrument puissant: la méthode comparative. Désormais, *fēc-* se présente non plus seulement dans le système des formes latines, mais entre dans le réseau des formes apparentées de langues congénères. En l'espèce, une donnée du grec ancien se révèle du plus haut intérêt: c'est l'aoriste hom. $\vartheta\eta\chi\epsilon$ « il posa, établit ». L'écart de sens entre *fēcit* et $\vartheta\eta\chi\epsilon$ ne fait pas de difficulté, car les conditions réelles de l'emploi dans l'une et l'autre langue fournissent les moyens d'une conciliation. Ainsi, lat. *fēc-* et gr. $\vartheta\eta\chi-$ se recouvrent parfaitement, c'est-à-dire continuent une seule et même forme préhistorique (**dhēk-*). Or, dans le système du grec, $\vartheta\eta\chi\epsilon$ se trouve en relation paradigmatique avec le présent $\tau\acute{\iota}\vartheta\eta\sigma\iota$ « il pose », analysable en $\tau\iota-+-\vartheta\eta-+-\sigma\iota$. On saisit donc une forme radicale $\vartheta\eta-$ plus élémentaire que $\vartheta\eta\chi-$. Du même coup, l'élément *-k-* de $\vartheta\eta\chi-$, *fēc-* se définit comme une consonne additionnelle. Et, de fait, le sanskrit n'en a pas trace, avec un aoriste *ádhāt*, un parfait *dadháu* et un présent *dádhati*, très proche de gr. $\tau\acute{\iota}\vartheta\eta\sigma\iota$. Au total, la comparaison permet un progrès dans l'analyse: De **dhēk-*, l'indo-européaniste descend jusqu'à **dhē-*, qui est une unité morphologique irréductible, une *racine*. Dans la théorie linguistique, la racine est une abstraction, une sorte de schéma opératoire. Actualisée dans une forme, c'est le *radical*. La connaissance de la racine et de ses différents « états » joue un rôle capital en grammaire comparée. La délimitation des derniers composants du signe n'est pas une fin en soi. Bien au contraire: analyser, décomposer, disséquer les formes, c'est se donner les moyens de rapprochements nouveaux. Qui s'avise de l'existence et des propriétés de la racine étend considérablement le champ de la comparaison. Aux correspondances de formes complètes s'ajoutent alors d'innombrables concordances partielles. Si lat. *vehit* « il transporte en char » recouvre parfaitement skr. *váhati* « il va en char » et si lat. *nouus* « nouveau » se superpose exactement à skr. *návaḥ* et à gr. *véφος*, en revanche l'identité formelle se limite à la racine dans lat. *tendō* vis-à-vis de gr. $\tau\epsilon\acute{\iota}\nu\omega$ « je tends » ou dans lat. *tepeō* « être chaud » en face de skr. *tápati* « brûler ». On pose donc, en indo-européen, d'une part des mots entiers, comme **wégheti* et **néwos*, de l'autre des racines, comme **ten-* et **tep-*. Le schème trilitère de ces unités est un trait constant de l'élément radical. Il n'y a pas de racines bilitères ou quadrilitères. Ainsi, la forme **dhē-*,

postulée tout à l'heure, se résout en **dheH-*; c'est-à-dire en une dentale sonore aspirée + une voyelle *e* + une consonne laryngale. Quant aux complexes quadrilitères du type **yeug-* (cf. gr. ζεύγυμι), ce sont en réalité des racines suffixées. On a donc **yéu-* + *-g-*. L'autonomie des composants ressort déjà de l'existence d'une structure inverse: **yu-ég-* (cf. skr. *yu-n-ák-ti*, avec infixé nasal). Au terme d'une démonstration géniale, E. Benveniste résumait ainsi ce balancement: «La racine fournit, avec un suffixe, deux thèmes alternants: I. racine pleine et tonique + suffixe zéro; II. racine zéro + suffixe plein et tonique»².

Aboutissement d'une analyse pénétrante, cette théorie marque le début d'un grand travail de synthèse. Des faits autrefois inclassables sortent de leur isolement et s'intègrent harmonieusement au système. Des anomalies disparaissent. Entre des formes en apparence hétérogènes des rapports naguère insoupçonnés deviennent manifestes. L'étude du système verbal se renouvelle. En particulier, l'Allemand Kl. Strunk met à profit la connaissance des thèmes II de racine et découvre la parenté étroite entre un type d'aoriste et les présents à nasale³. L'Américain C. Watkins donne ensuite une histoire magistrale du verbe indo-européen⁴. Enfin, une présentation nouvelle de la morphologie nominale est actuellement en préparation. En un mot, la construction de la grammaire indo-européenne se fait sur nouveaux frais. Sous l'impulsion de la linguistique générale, l'esprit de synthèse prend aujourd'hui le pas sur le souci d'analyse. Au fur et à mesure des progrès de la recherche typologique, le comparatiste se préoccupe davantage de la cohérence interne de ses reconstructions. L'indo-européen n'échappe pas aux conditions générales des systèmes de signes. C'est pourquoi, si le produit d'une restitution se trouve en désaccord avec les universaux des langues naturelles, une remise en question s'impose. On reprend alors les données et on s'interroge sur le bien-fondé de certains rapprochements. Cela explique l'épanouissement de toute une réflexion méthodologique. Des théoriciens établissent avec rigueur les principes de la comparaison. Une correspondance probante, c'est-à-dire révélatrice d'une parenté, suppose: 1° au niveau formel l'existence de rapports constants, vérifiables à l'intérieur d'une série de termes; 2° au niveau du sens une situation de synonymie ou du moins de compatibilité. Au stade actuel du développement de la discipline, la vérification du premier point fait rarement difficulté. En revanche, l'examen de la deuxième condition pose fréquemment des problèmes délicats. Quand les sens de termes présumés correspondants ne concordent pas exactement, comment distinguer dans la pratique entre les différences non signifi-

catives et les écarts dirimants? Cette aporie retient de plus en plus l'attention des chercheurs. L'exigence de l'adéquation sémantique des éléments d'une comparaison ne le cède pas à la nécessité de la régularité formelle⁵. En chaque cas, les membres d'une correspondance font donc l'objet d'une définition précise. La méthode est, une fois de plus, synthétique: on rassemble tous les emplois des mots en litige, on en relève toutes les constructions syntaxiques et, sur la base de l'ensemble de ces faits, on se prononce finalement pour ou contre une relation de parenté. De la justesse de l'appréciation dépend en dernier ressort la restitution du lexique indo-européen. Comme le vocabulaire de n'importe quelle



	Dupont	Durand	Autre
Homme	△	▲	
Femme	○	●	⊙
= Alliance		— Filiation	

langue, ce lexique présente une organisation propre. Par le jeu des associations se constituent des groupes d'unités solidaires. Il y a, par exemple, un système des termes de parenté. Sa configuration reflète dans une large mesure la structure de la famille et de la société. L'étude linguistique s'appuie donc en ce domaine sur l'anthropologie et l'ethnologie. Ainsi, par référence à l'institution du mariage entre cousins croisés, bien connue des ethnologues, l'indo-européaniste explique des relations formelles en apparence insolites. A titre d'illustration, la désignation latine de l'oncle maternel par un dérivé de *auus* « grand-père », à savoir le diminutif *auunculus*, se justifie désormais sans peine. En effet, dans le type de société hérité par les Romains, mon grand-père paternel est en même temps mon grand-oncle maternel (voir le schéma ci-dessus, tiré d'E. Benveniste, *Le vocabulaire des institutions indo-européennes* I, Paris, 1969, 228). Ce n'est donc pas sans raison que l'oncle est appelé « petit grand-père ».

En définitive, on le voit, l'effort de synthèse au plan du lexique indo-européen ouvre des perspectives non seulement linguistiques, mais extra-linguistiques aussi bien. Le vocabulaire le plus ancien comporte, en effet, un enseignement irremplaçable sur le peuple indo-européen. Ces lointains ancêtres possédaient déjà une riche civilisation, avec une puissante organisation sociale, une force guerrière entreprenante, des institutions religieuses et, probablement, une poésie épique de tradition orale⁶. L'image globale de cette culture procède encore d'une synthèse, par reconstruction au second degré en quelque sorte. On a maintenant une idée du champ de la grammaire comparée et de sa place au sein des sciences humaines. Le domaine de ces études est très vaste et suscite à tout moment la collaboration interdisciplinaire. Mais, pour l'accomplissement de sa tâche immense, le comparatiste est servi par une méthode sûre et éprouvée, qui conjugue avec bonheur l'analyse et la synthèse.

NOTES

¹ Über ein Gesetz der indogermanischen Wortstellung: *Indogermanische Forschungen* I, 1892, 333-436.

² *Origines de la formation des noms en indo-européen*, Paris 1935, 170.

³ *Nasalpräsentien und Aoriste*, Heidelberg 1967.

⁴ *Indogermanische Grammatik III 1: Geschichte der indogermanischen Verbal-flexion*, Heidelberg 1969.

⁵ Voir E. Benveniste, Problèmes sémantiques de la reconstruction: *Word* 10, 1954, 251-264 (= *Problèmes de linguistique générale* 1, Paris 1966, 289-307).

⁶ Cf. R. Schmitt, *Dichtung und Dichtersprache in indogermanischer Zeit*, Wiesbaden 1967.