



**HAL**  
open science

# Les Indo-européens et leur langue : reconstruction d'une reconstruction

Jean-Philippe Watbled

► **To cite this version:**

Jean-Philippe Watbled. Les Indo-européens et leur langue : reconstruction d'une reconstruction. Travaux & documents, 2015, 48, pp.183–200. hal-01501124

**HAL Id: hal-01501124**

**<https://hal.univ-reunion.fr/hal-01501124>**

Submitted on 29 Sep 2017

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

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

# Les Indo-européens et leur langue : reconstruction d'une reconstruction

---

JEAN-PHILIPPE WATBLED  
PROFESSEUR, MEMBRE DU LCF  
UNIVERSITÉ DE LA RÉUNION

L'objectif de cet article est de faire un tour d'horizon des recherches dans le domaine indo-européen et de fournir quelques éléments de réflexion sur un domaine complexe qui demeure aujourd'hui et sans doute pour longtemps encore l'objet de vifs débats. On abordera les problèmes proprement linguistiques<sup>1</sup>, mais aussi celui de la localisation du foyer originel du peuple indo-européen présumé, ce qui sera l'occasion d'évoquer les critiques et les approches sceptiques à l'égard de l'hypothèse indo-européenne classique, pour terminer par une rapide présentation d'autres hypothèses plus ambitieuses.

Commençons par poser deux questions en guise de préambule : Qu'est-ce que l'indo-européen ? Qui étaient les Indo-européens ? Les Indo-européens, peuple dont on n'a aucune trace, étaient les locuteurs de l'indo-européen, langue dont on n'a aucune trace. Voilà ce que nous dit ce que j'appellerai ici, pour les besoins de la cause, la théorie standard. Ce constat est d'autant plus étonnant et remarquable que les réflexions ne portent pas exclusivement sur la langue : ainsi Georges Dumézil, par exemple, s'est intéressé non seulement aux langues, mais aussi aux religions et à la mythologie des Indo-européens<sup>2</sup>, et a postulé sa théorie bien connue de la trifonctionnalité (fonction sacerdotale ; fonction guerrière ; fonction productrice), dont certains se demandent si elle est distinctive des Indo-européens ou si elle correspond à des universaux. Remarquons que les recherches sur la mythologie indo-européenne n'existeraient pas sans l'hypothèse d'un peuple indo-européen, qui elle-même découle de l'hypothèse linguistique d'une langue source à l'origine de toutes les langues dites indo-européennes : c'est donc *in fine* la question de la langue qui prime.

Cette idée d'une langue très ancienne est née de l'observation de correspondances relativement régulières entre des groupes de langues distincts. Comme l'écrit Pierre Monteil, « la distance géographique ou chronologique

---

<sup>1</sup> Sur ce plan, même si l'on ne se limitera pas à la phonologie, l'accent sera mis principalement sur cette composante du langage dont il importe de rappeler le caractère fondamental, car finalement tout passe d'abord et avant tout par la parole.

<sup>2</sup> Georges Dumézil, *Mythes et dieux des Indo-européens*. Paris : Flammarion, 1992.

excluant entre ces langues l'emprunt, la spécificité des notions étudiées excluant le hasard, seule subsiste comme explication raisonnable une origine commune »<sup>3</sup>.

Un certain nombre de langues anciennes formeraient donc une famille. Il faut ajouter que l'indo-européen n'est que l'une des familles postulées pour la planète ; au même titre que l'indo-européen, on a, par exemple, sans viser aucunement l'exhaustivité, les familles suivantes, chacune comprenant des groupes, des sous-groupes, etc. : la famille dravidienne (langues du sud de l'Inde), la famille ouralique (incluant le finno-ougrien), la famille afro-asiatique (incluant le sémitique), la famille australienne, et bien d'autres encore. L'indo-européen mérite néanmoins une observation : son destin a été remarquable, puisque ce sont des locuteurs de langues indo-européennes qui ont colonisé une grande partie du monde.

Les recherches indo-européennes commencent au XVI<sup>e</sup> siècle. Des missionnaires jésuites ont alors constaté une ressemblance entre le sanskrit et les langues classiques de l'Europe, le grec et le latin. En 1768, le jésuite Gaston-Laurent Coeurdoux (1691-1777) présentait à l'Académie française des éléments montrant la parenté entre ces langues : au vu des correspondances, le hasard ne pouvait guère être invoqué. En 1786, Sir William Jones (1746-1794), linguiste et orientaliste, vantait dans un discours devant la Société Asiatique du Bengale la « merveilleuse structure » du sanskrit, qu'il considérait comme supérieur au grec et au latin à bien des égards. Mais ce qui comptait le plus pour William Jones, c'étaient les affinités entre ces trois langues, qui faisaient penser avec certitude à une source commune, qui « peut-être n'existait plus » ; et il ajoutait avec une remarquable intuition que le gothique, le celtique et le vieux perse avaient probablement la même origine. William Jones avait déjà bien vu que les correspondances concernaient non seulement le lexique, mais aussi la grammaire. Or on sait que les langues sont « emprunteuses » surtout en matière de lexique, alors que les composantes phonétique et grammaticale préservent largement leur autonomie : les correspondances dans ces deux dernières composantes sont donc cruciales, le lexique présentant des risques d'erreurs ; il n'est en effet pas toujours facile de savoir si une ressemblance lexicale entre plusieurs langues s'explique par la parenté historique ou par l'emprunt. En tout cas, l'aventure des études indo-européennes avait bel et bien commencé.

La protolangue a été baptisée « indo-européen » par Thomas Young en 1813. Parmi les premiers grands indo-européanistes, on peut citer Rasmus Rask (1787-1832), Friedrich Schlegel (1772-1829), Jakob Grimm (1785-1863), le grand comparatiste Franz Bopp (1791-1867), à qui l'on doit la standardisation du terme « indo-européen », August Friedrich Pott (1802-1887), qui a publié (1833-1836) le

---

<sup>3</sup> Pierre Monteil, *Éléments de phonétique et de morphologie du latin*. Paris : Nathan, 1986, p. 14.

fruit de ses recherches étymologiques sur la sphère de ce qu'il a appelé les langues « indo-germaniques », terme qui devait alors concurrencer l'appellation d'indo-européen, et qui a encore cours en Allemagne, August Schleicher (1821-1868), Karl Brugmann (1849-1919), etc. En France, Antoine Meillet (1866-1936) jouera un rôle décisif et sera considéré comme un maître des études indo-européennes, tout comme Émile Benveniste (1902-1976), auteur du célèbre *Vocabulaire des institutions indo-européennes*<sup>4</sup>, sans oublier André Martinet (1908-1999), l'une des grandes figures du structuralisme fonctionnaliste<sup>5</sup>. On ne peut évidemment rendre justice à tout le monde, d'autant plus que les études indo-européennes n'ont jamais cessé de s'étendre et de se renouveler.

À la suite du travail des pionniers, on a su que, selon la théorie standard, la famille indo-européenne était structurée en plusieurs groupes majeurs<sup>6</sup> : le celtique, l'italique (avec entre autres le latin, son représentant le plus célèbre), le germanique, le balte et le slave (parfois regroupés en balto-slave), l'albanais, le grec (avec tous ses « dialectes »), l'arménien, l'anatolien, et l'indo-iranien, lui-même constitué de l'indo-aryen<sup>7</sup> et de l'iranien.

Le groupe indo-iranien comprend, entre autres, deux langues sacrées de grand prestige, et qui sont linguistiquement fortement reliées : l'avestique, langue iranienne de l'Avesta, ensemble des textes sacrés de la religion mazdéenne, et le sanskrit, langue indo-aryenne des Veda<sup>8</sup>.

Le mot « sanskrit » vient de *saṃskṛtam* dérivé du participe passé *saṃskṛtaḥ* du verbe *saṃ-s-kr-*, « réunir, arranger, consacrer, purifier » (*saṃ*, « ensemble » ; *kr-*, « faire »), d'où « langue raffinée, sacrée »<sup>9</sup>. Les spécialistes du

<sup>4</sup> Émile Benveniste, *Vocabulaire des institutions indo-européennes*, deux volumes. Paris : Editions de Minuit, 1969.

<sup>5</sup> Pour un bon résumé des thèses d'André Martinet sur l'indo-européen, on peut se reporter à son ouvrage *Des steppes aux océans*. Paris : Payot, 1986/1994.

<sup>6</sup> Je laisse ici de côté l'intégration de groupes mineurs ou de langues mineures.

<sup>7</sup> Le mot sanskrit *ārya-*, issu d'un terme proto-indo-iranien par lequel les Indo-iraniens se désignaient eux-mêmes, se définit ainsi : « Aryen, hab. de l'Āryāvarta ; homme des trois premières castes ; homme honoré (aïeul, frère aîné, etc. ; titre honorifique) ; [...] noble, aryen, honorable [...] » (Nadine Stchoupak, Luigia Nitti et Louis Renou, *Dictionnaire sanskrit-français*. Paris : Librairie d'Amérique et d'Orient, Jean Maisonneuve successeur, sixième tirage, 1987, p. 123). La notion dominante semble être celle d'honneur et de noblesse. Le mot *āryāvarta*, composé de *ārya-* et de *āvarta* (« tourbillon », « mèche de cheveux », mais aussi « endroit très peuplé »), signifie « terre des Aryens » et désigne le nord de l'Inde dans la littérature classique en sanskrit.

<sup>8</sup> Le recueil le plus connu est le Rig-Veda (ou *Rgveda*).

<sup>9</sup> Nadine Stchoupak, Luigia Nitti et Louis Renou, *op. cit.*, p. 758-759. Le terme *prākṛit* (sanskrit *prākṛta*) désigne « tout dialecte apparenté au sanskrit » (*op. cit.*, p. 492), le sens général étant

sanskrit ont pu profiter du travail de Pāṇini<sup>10</sup>, le grand linguiste indien de l'Antiquité, dont la grammaire, intitulée *Aṣṭādhyāyī*<sup>11</sup>, est de très loin l'exemple le plus remarquable de traité de linguistique descriptive et théorique de l'Antiquité, n'ayant rien à envier à notre meilleure linguistique contemporaine.

Concernant le grec, la forme la plus ancienne que nous en connaissions est le mycénien. Les inscriptions dans ce dialecte ont été préservées sur des tablettes d'argile et des vases de céramique en Crète (notamment à Gnosso, 1400-1350 avant notre ère) et en Grèce continentale (notamment à Pylos, 1200 avant notre ère). L'écriture de cette civilisation mycénienne est ce qu'on appelle le linéaire B, de type syllabique, qui a été décodé par Michael Ventris (1922-1956) en 1952. Celui-ci a ensuite poursuivi son travail en collaboration avec John Chadwick (1920-1998) : ils ont publié ensemble les *Documents in Mycenaean Greek* en 1955<sup>12</sup>. Notons que plusieurs siècles s'écouleront entre les inscriptions en linéaire B et les textes postérieurs d'une certaine ampleur dont nous disposons, à savoir les épopées homériques.

Mais notre intérêt ne doit en aucun cas se limiter aux langues les plus prestigieuses, et il importe de savoir que les dernières découvertes dues aux recherches archéologiques dans le domaine indo-européen sont d'une part le tokharien, subdivisé en deux dialectes dits « A » et « B », dont on a trouvé les traces à la fin du XIX<sup>e</sup> siècle lors d'expéditions occidentales dans le Xinjiang, province du nord-ouest de la Chine, et d'autre part le hittite<sup>13</sup>, du groupe anatolien, pour lequel les premiers documents, en graphie cunéiforme akkadienne et en graphie hiéroglyphique, ont été mis au jour au début du XX<sup>e</sup> siècle seulement. Le hittite a été reconnu comme faisant partie de l'indo-européen en 1915. Avec cette langue, nous tenons les traces les plus anciennes d'une langue indo-européenne (XVIII<sup>e</sup>/XVII<sup>e</sup> siècles avant notre ère).

À présent, avant de discuter de la validité de l'hypothèse indo-européenne, quelques mots d'explication sont nécessaires sur la méthode comparative en linguistique, qui a justement permis de postuler l'existence d'une protolangue. Il

« original, naturel, primitif [...] », *ibid.* Les langues indo-aryennes modernes dérivent des *prākritis*.

<sup>10</sup> On ne s'accorde pas sur la période durant laquelle a vécu Pāṇini (en tout cas, entre le VII<sup>e</sup> et le IV<sup>e</sup> siècle avant notre ère).

<sup>11</sup> Littéralement : « Huit chapitres » : *aṣṭa*, « huit », *adhyāya*, « lecture, temps de lecture, chapitre » (Nadine Stchoupak, Luigia Nitti et Louis Renou, *op. cit.*, p. 25).

<sup>12</sup> John Chadwick et Michael Ventris, *Documents in Mycenaean Greek*. Cambridge: Cambridge University Press (2<sup>e</sup> éd.), 1974.

<sup>13</sup> Calvert Watkins, « Hittite », in Roger D. Woodard (dir.), *The Ancient Languages of Asia Minor*. Cambridge: Cambridge University Press, 2008, p. 6-31, et Theo van den Hout, *The Elements of Hittite*. Cambridge: Cambridge University Press, 2011.

existe *a priori* plusieurs sources de ressemblances entre les langues<sup>14</sup>. La première est le hasard : deux mots de deux langues différentes, et de signification apparentée, peuvent se ressembler phonétiquement, par les aléas de deux histoires séparées ; la seconde est le contact de langues, qui est plus exactement un contact entre des communautés linguistiques différentes, susceptible d'installer dans certains cas des affinités dites aréales ; la troisième est tout simplement l'emprunt, lexical principalement, qui suppose aussi une forme de contact ; la quatrième, enfin, est l'existence d'universaux du langage, objet d'étude des spécialistes de typologie des langues : les universaux sont souvent responsables d'analogies entre des langues sans filiation ni contact. Toutefois, il existe des situations où ces différentes pistes doivent être écartées. Prenons l'exemple des termes « père » et « mère » dans trois langues classiques<sup>15</sup> :

<u>sanskrit</u>	<u>grec</u>	<u>latin</u>
<i>pitár-</i>	<i>patḗr</i>	<i>páter</i>
<i>mātár-</i>	<i>mētḗr</i>	<i>māter</i>

Précisons qu'il ne s'agit là que d'exemples représentatifs<sup>16</sup> parmi bien d'autres, sachant que les correspondances ne sont pas toujours aussi fortes ou générales, ou pas aussi systématiques. Néanmoins, elles sont assez nombreuses pour que l'on écarte le hasard ou les contacts de langues différentes au départ. Dans de tels cas, il est légitime de poser une source commune. Il faut également disposer de la méthodologie appropriée pour déterminer la nature de la source en question. Par ailleurs, ce qui est à reconstruire est un ensemble organisé et systématique pour chaque composante de la langue – phonologie, lexique, grammaire – et non des éléments isolés<sup>17</sup>.

<sup>14</sup> Benjamin W. Fortson IV, *Indo-European Language and Culture. An Introduction*. Hoboken : Wiley-Blackwell (2<sup>e</sup> éd.) 2010, p. 1-3.

<sup>15</sup> James Clackson, *Indo-European Linguistics. An Introduction*. Cambridge: Cambridge University Press, 2007, p. 202. Le macron sur une voyelle signale qu'elle est longue, et l'accent porte sur la voyelle de la syllabe dominante. Pour le sanskrit, *pitár-* et *mātár-* sont les bases flexionnelles, alors que pour le grec et le latin, les formes données sont celles du nominatif singulier.

<sup>16</sup> Pour la reconstruction des sources indo-européennes de ces formes, voir plus loin. Notons qu'en grec, la forme *mētḗr* est celle du grec ionien-attique, alors que le dialecte dorien avait une forme ancienne *mātḗr*, plus proche du latin et du sanskrit : voir Michel Lejeune, *Phonétique historique du mycénien et du grec ancien*. Paris : Klincksieck, 1972, p. 191, p. 234-235.

<sup>17</sup> Toute méthode de reconstruction se heurte inévitablement à la question de la variation : il est déjà très difficile d'envisager de reconstruire un système hypothétique, il est encore plus difficile de simplement envisager de reconstruire la variation ! Dans cet esprit, toute reconstruction est une abstraction purement théorique.

En dépit de son caractère probant et de ses résultats, il arrive que des chercheurs doutent de la validité de la méthode de reconstruction (voir *infra*). Or certains événements de l'histoire des études indo-européennes méritent d'être signalés à l'attention des sceptiques. Je mentionnerai ce qui est sans doute la plus remarquable de ces avancées. Ferdinand de Saussure (1857-1913), dans son monumental *Mémoire sur le système primitif des voyelles dans les langues indo-européennes* (écrit à l'âge de 21 ans, en 1878), a montré que certaines irrégularités morphologiques observées dans les langues classiques trouvent leur solution si l'on postule pour la protolangue des éléments phonologiques disparus dans les données observables en latin, grec, sanskrit, etc. Ainsi, il a avancé l'hypothèse que toute voyelle longue alternant dans les paradigmes flexionnels avec une voyelle brève a en réalité comme source indo-européenne une voyelle brève associée à ce qu'il a appelé un « coefficient sonantique », et pour rendre compte des données, il a posé trois de ces « coefficients » différents. Dans l'esprit de Saussure, il ne s'agissait là que de formules algébriques censées avoir une valeur explicative.

Mais en 1927, le linguiste Jerzy Kurylowicz a démontré que les « coefficients sonantiques » en question se retrouvent en partie en hittite, dans des documents découverts ultérieurement, au début du XX<sup>e</sup> siècle (voir *supra*), sous la forme d'une consonne transcrite simple ou double, ce qui laisse supposer qu'il s'agissait de deux phonèmes proches, mais distincts. Le hittite est ainsi la seule langue indo-européenne ancienne attestée à avoir préservé au moins un sous-ensemble des consonnes correspondant aux « coefficients sonantiques » de Saussure. Ces consonnes ont été baptisées « laryngales »<sup>18</sup>, à tort dans la mesure où l'on ne connaît pas leur nature phonétique exacte : on peut simplement supposer qu'elles étaient articulées à un niveau profond du tractus vocal (uvulaires, pharyngales ou glottales).

Dans les approches les plus courantes, on postule trois « laryngales », représentées par les symboles  $H_1$ ,  $H_2$ ,  $H_3$ . Deux de ces trois consonnes ont manifestement exercé un pouvoir « colorant » sur les voyelles qui les jouxtaient : pour  $H_2$  un pouvoir « colorant » de type *a*, et pour  $H_3$  un pouvoir « colorant » de type *o*. Les « laryngales » colorantes étaient possiblement des fricatives uvulaires<sup>19</sup> : par exemple \* $\chi$  pour la coloration *a* ( $H_2$ ) et \* $\chi^w$  pour la coloration *o* ( $H_3$ )<sup>20</sup>.

<sup>18</sup> Cette appellation est due à Hermann Möller en 1911.

<sup>19</sup> À ce sujet, voir André Martinet, *Des steppes aux océans*, *op. cit.*, ch. IX.

<sup>20</sup> Le symbole  $\chi$  dénote un son proche de celui du *ch* allemand de *acht* ou du *j* espagnol. Le symbole <sup>w</sup> dénote quant à lui une articulation secondaire de type labial (cf. le *w* de *watt*). Conformément à l'usage établi, l'astérisque (\*) précède les formes reconstruites, donc non attestées.

Les études qui s'ensuivront sur les mots hittites comportant ces « laryngales », ainsi que les comparaisons avec les autres langues anciennes, confirmeront de manière éclatante l'hypothèse saussurienne. Depuis, l'accord ne se fait toujours pas sur le nombre et la nature exacte des « laryngales » de l'indo-européen, mais ceux qui rejettent l'hypothèse elle-même sont nettement minoritaires.

Pour les exemples des mots « père » et « mère » donnés *supra*, les formes reconstruites sont *\*pH<sub>2</sub>ter-* et *\*meH<sub>2</sub>ter-*. Le produit des « laryngales » interconsonantiques, comme dans *\*pH<sub>2</sub>ter-*, est régulièrement vocalique, le plus souvent *a* (cf. *páter* en latin), mais *i* en sanskrit (cf. *pitár-*). En grec, *H<sub>1</sub>*, *H<sub>2</sub>*, *H<sub>3</sub>* ont donné *e*, *a*, *o*, respectivement<sup>21</sup>. Les « laryngales » se sont par ailleurs effacées au contact d'une voyelle, entraînant un allongement compensatoire de celle-ci en contexte final ou préconsonantique, comme dans *\*meH<sub>2</sub>ter-* > sanskrit *mātár-*, grec *mētēr*, latin *māter*.

Malgré les exploits intellectuels – comme celui de Saussure, qui vient d'être exposé – et leurs conséquences positives, il n'y a pas unanimité sur la reconstruction du système phonologique de la protolangue, et il n'y en aura sans doute jamais, en raison même de la nature de l'entreprise. Il importe donc de bien distinguer à cet égard l'hypothèse de la source proto-indo-européenne – dont on peut à mon avis être certain – et les structures de la langue en question, que l'on a du mal à déterminer.

Ainsi pour les consonnes plosives, la méthode de reconstruction standard donne les classes suivantes<sup>22</sup> :

	labiales	dentales	palatales	vélaires	labio-vélaires
non voisées	*p	*t	*ḳ	*k	*k <sup>w</sup>
voisées	(*b)	*d	*g̣	*g	*g <sup>w</sup>
voisées aspirées	*b <sup>h</sup>	*d <sup>h</sup>	*g <sup>h</sup>	*g <sup>h</sup>	*g <sup>wh</sup>

Cette reconstruction pose différents problèmes, que nous allons examiner successivement. Le premier est le fait que l'on ne retrouve jamais les trois classes de dorsales (palatales, vélaires, labio-vélaires) réunies dans une même langue indo-

<sup>21</sup> Benjamin W. Fortson IV, *op. cit.*, p. 62-63.

<sup>22</sup> Voir par exemple Benjamin W. Fortson IV, *op. cit.*, p. 56, et James Clackson, *op. cit.*, p. 34. Les palatales, vélaires et labio-vélaires constituent les trois sous-classes de consonnes dorsales. Précisons qu'il est impossible de connaître la nature phonétique exacte des palatales du proto-indo-européen. D'une manière générale, les palatales sont articulées aux environs du lieu de la voyelle *i* (= palais dur), alors qu'une articulation vélaire est postérieure et correspond au lieu de la voyelle du français *ou* (= voile du palais, dit aussi palais mou). Pour les aspirées, l'articulation proprement dite est suivie d'un souffle glottal, combiné avec le voisement (= vibrations des plis vocaux) dans le cas des aspirées voisées.



européenne attestée<sup>23</sup>. Les trois rangées de dorsales sont en effet réduites à deux dans pratiquement toutes les langues de la famille, les palatales ou les labio-vélares (selon les groupes de langues) ayant fusionné avec les vélares. Dans le premier type de développement, la classe des vélares (issues des palatales et des vélares originelles) s'oppose à celle des labio-vélares, et dans le second la classe des vélares (issues des vélares et des labio-vélares originelles) s'oppose aux produits des palatales, c'est-à-dire à des chuintantes ou des sifflantes (affriquées<sup>24</sup> ou non). Le premier type de développement a été appelé *centum* et le second *satem*, sur la base de la forme prise par le terme signifiant « cent » en latin et en avestique respectivement<sup>25</sup>. On distingue alors des langues de type *centum*, comme le latin classique qui possède des vélares et des labio-vélares<sup>26</sup>, mais pas de palatales, et des langues de type *satem*, comme le sanskrit qui possède des chuintantes palato-alvéolaires issues de \*k̑ et \*g̑, ainsi que des vélares, mais pas de labio-vélares<sup>27</sup>.

Le fait que la coexistence des trois classes de dorsales ne soit pas attestée a amené certains, à la suite d'Antoine Meillet<sup>28</sup>, à ne postuler que deux classes de dorsales pour le proto-indo-européen, les vélares et les labio-vélares : les vélares se seraient d'abord palatalisées, puis assibilées<sup>29</sup> avant des vocoïdes<sup>30</sup> antérieures,

---

<sup>23</sup> Voir cependant le cas du louvite, évoqué plus loin.

<sup>24</sup> Les affriquées sont des consonnes uniques qui commencent comme des plosives et se terminent par un relâchement fricatif, comme le *ch* de l'espagnol *mucho*, par exemple.

<sup>25</sup> On sait aujourd'hui que cette distinction entre *centum* et *satem* ne correspond pas à des ensembles dialectaux homogènes, surtout pour ce qui concerne le type *centum*.

<sup>26</sup> Le grec ancien, de type *centum*, a « perdu » également les labio-vélares, les transformant en labiales, dentales ou vélares simples selon le contexte (Michel Lejeune, *op. cit.*, p. 35 et p. 43-53). Avant le déchiffrement du linéaire B (voir *supra*), on avait postulé que le protogrec, source non attestée des dialectes des époques archaïque et classique, possédait les labio-vélares, qui auraient été transformées ensuite par des processus phonétiques. Or, grâce au déchiffrement du linéaire B, on a découvert dans les documents mycéniens un signe correspondant spécifiquement aux labio-vélares, ce qui révélait clairement que le système du mycénien comprenait encore cette classe de consonnes issues du proto-indo-européen.

<sup>27</sup> La situation du sanskrit est complexe : la chuintante non voisée issue de \*k̑ est une fricative (cf. *ch* du français *chou*), alors que la chuintante voisée issue de \*g̑ est une affriquée (cf. *j* de l'anglais *jam*) ; en outre, les vélares résultant de la fusion des vélares et des labio-vélares originelles du proto-indo-européen se sont elles-mêmes palatalisées et transformées en chuintantes affriquées (cf. *ch* et *j* de l'anglais *church* et *jam*) avant les vocoïdes antérieures.

<sup>28</sup> Antoine Meillet, *De quelques difficultés de la théorie des gutturales indo-européennes*, extrait des *Mémoires de la société linguistique de Paris*, t. VIII, 1894 (Source gallica.bnf.fr / Bibliothèque nationale de France).

<sup>29</sup> C'est-à-dire : transformées en sifflantes ou chuintantes.

<sup>30</sup> On regroupe sous le terme de *vocoïdes* les voyelles et les semi-voyelles.

les « palatales » étant alors à considérer comme des produits secondaires, postérieurs à la protolangue<sup>31</sup>.

La théorie des deux classes de dorsales a toutefois été mise à mal par le travail de H. Craig Melchert sur le louvite<sup>32</sup> : les données issues de cette langue anatolienne semblent en effet aller dans le sens d'une distinction originelle entre palatales, vélares et labio-vélares car, même si c'est en proportion limitée, les produits des trois rangées de dorsales y sont demeurés distincts<sup>33</sup>.

Mais la théorie des trois classes ne va pas sans inconvénient non plus, car dans un tel système, on s'attendrait à ce que les palatales, en principe marquées<sup>34</sup> et par conséquent plus complexes que les vélares, soient moins fréquentes que ces dernières : or c'est l'inverse que l'on constate dans les reconstructions. Par ailleurs, s'il est facile d'expliquer que dans les langues *satem*, les labio-vélares ont perdu leur articulation secondaire pour fusionner avec les vélares simples (\*k<sup>w</sup> > \*k, etc.), il est plus difficile de concevoir que dans les langues *centum*, les palatales se soient transformées en vélares (\*k̑ > \*k, etc.) pour fusionner avec les vélares originelles : tout spécialiste de linguistique historique et comparée sait en effet que si les vélares se palatalisent fréquemment, le processus inverse – le recul des palatales dans la région vélaire – est rarissime.

Une solution possible à ce double problème, dans la ligne de ce qui a été proposé par Martin E. Huld<sup>35</sup>, consiste à supposer que l'on avait effectivement trois classes de dorsales en proto-indo-européen, mais qu'il ne s'agissait pas des points d'articulation envisagés habituellement : là où la théorie standard postule des palatales, on aurait eu en fait des vélares, et là où elle postule des vélares, on

---

<sup>31</sup> Voir par exemple Winfred P. Lehmann, *Theoretical Bases of Indo-European Linguistics*. London & New York: Routledge, p. 100-101.

<sup>32</sup> H. Craig Melchert, « *Proto-Indo-European velars in Luvian* », in Calvert Watkins (dir.), *Studies in Memory of Warren Cowgill (1925-1985)*. Berlin : de Gruyter, 1987, p. 182-204. Voir aussi les remarques de Benjamin W. Fortson IV sur le louvite (*op. cit.*, p. 187).

<sup>33</sup> Les observations de Don Ringe sur les contraintes concernant les lieux d'articulation dans les racines indo-européennes viennent également à l'appui de l'hypothèse de trois classes de dorsales : voir Don Ringe, *From Proto-Indo-European to Proto-Germanic, A Linguistic History of English*, vol. I. Oxford: Oxford University Press, 2006, p. 7.

<sup>34</sup> Une unité ou une série d'unités est dite marquée si elle est plus complexe qu'une autre avec laquelle elle entretient une relation d'opposition.

<sup>35</sup> Martin E. Huld, « *Satəm, Centum, and Hokum* », in Douglas Q. Adams (dir.), *Festschrift for Eric P. Hamp, Volume I, Journal of Indo-European Studies Monograph Series 23*. Washington, D.C.: Institute for the Study of Man, 1997, p. 115-138.

aurait eu des uvulaires. Cela nous donne le tableau suivant, distinct du précédent<sup>36</sup> :

	labiales	dentales	vélaires	uvulaires	uvulaires labialisées
non voisées	*p	*t	*k	*q	*q <sup>w</sup>
voisées	(*b)	*d	*g	*G	*G <sup>w</sup>
voisées	*b <sup>h</sup>	*d <sup>h</sup>	*g <sup>h</sup>	*G <sup>h</sup>	*G <sup>wh</sup>
aspirées					

Dans ce cadre, les vélaires ne sont pas marquées, alors que les uvulaires, nettement plus difficiles à articuler, le sont, ce qui règle le premier problème évoqué plus haut. Dans toutes les branches de la famille indo-européenne, les uvulaires voient ensuite leur lieu d'articulation avancer au point vélaire : \*q > \*k, \*G > \*g, \*G<sup>h</sup> > \*g<sup>h</sup>, \*q<sup>w</sup> > \*k<sup>w</sup>, \*G<sup>w</sup> > \*g<sup>w</sup>, \*G<sup>hw</sup> > \*g<sup>hw</sup>. La différence essentielle entre les langues *centum* et les langues *satem* serait alors que dans les premières, les vélaires originelles ne changent pas, ce qui entraîne que les uvulaires simples se confondent avec elles, alors que dans les secondes, les vélaires originelles se palatalisent sous la pression paradigmatique de l'avancée des uvulaires simples<sup>37</sup> : \*k > \*k̠, \*g > \*g̠, \*g<sup>h</sup> > \*g̠<sup>h</sup>.

On observe que dans cette approche, on considère que les palatales des langues *satem* résultent de l'avancée de vélaires originelles, tout comme dans la théorie des deux classes de dorsales. Dans les langues *satem*, la palatalisation des vélaires est suivie de la perte de l'articulation secondaire des labio-vélaires, qui se confondent ainsi finalement avec les nouvelles vélaires issues des uvulaires simples : \*k<sup>w</sup> > \*k, \*g<sup>w</sup> > \*g, \*g<sup>wh</sup> > \*g<sup>h</sup>.

Cette hypothèse présente l'avantage de rendre compte de l'existence des trois classes de dorsales (voir la question du louvite *supra*), tout en n'ayant pas à postuler que les palatales des langues *centum* comme le latin auraient subi un processus peu plausible de dépalatalisation. Ensuite, elle facilite la reconstruction de la source proto-indo-européenne de cognats tels que le grec *ákṃōn* (« enclume ») et le sanskrit *ásman-* (« pierre ») : en raison de la chuintante palato-alvéolaire *ʃ* du sanskrit, il faudrait dans la variante du cadre standard avec trois classes de dorsales postuler une source *\*H<sub>2</sub>ǵk̠mōn* (« pierre ») avec une palatale (\*k̠) pour rendre compte du *ʃ* du sanskrit, et dans la variante avec deux classes de dorsales, il faudrait partir de *\*H<sub>2</sub>ǵk̠mōn*, mais on se demanderait bien

<sup>36</sup> Les symboles q et G dénotent les plosives uvulaires, qui ont le même mode d'articulation que k et g, mais nettement plus en arrière, au niveau de la luette.

<sup>37</sup> Cette palatalisation a sans aucun doute été favorisée dans le contexte de vocoïdes antérieures.

alors ce qui aurait pu conditionner la palatalisation de \*k dans un tel contexte phonétique dans les langues *satem* telles que le sanskrit. Dans l'hypothèse défendue ici, on partira de \**H<sub>2</sub>ek-mōn*, et le \*k subit la palatalisation en raison non pas du contexte phonétique, peu propice, mais de la pression paradigmatique signalée *supra*, ce qui est très différent.

L'autre grand problème lié au système des plosives concerne les traits de phonation. La théorie standard postule trois séries : une série de plosives non voisées, une série de plosives voisées simples et une série de plosives voisées aspirées. Or du point de vue typologique, un tel système est peu plausible : en effet, on imagine difficilement une série de voisées aspirées sans des partenaires non-voisées aspirées (\*p<sup>h</sup>, \*t<sup>h</sup>, etc.)<sup>38</sup>, normalement moins marquées. Pour remédier à cela, on a proposé l'hypothèse glottalique, selon laquelle les consonnes de la série voisée de la théorie standard auraient été en réalité des éjectives, c'est-à-dire des consonnes à initiation glottalique dans la production desquelles l'air est comprimé entre la glotte verrouillée et le point d'articulation supraglottal. Il existe différentes variantes de la théorie glottalique<sup>39</sup>. Personnellement, je suis tenté de postuler pour les plosives le système suivant, qui intègre en outre l'hypothèse « uvulaire » exposée plus haut :

	labiales	dentales	vélaires	uvulaires	uvulaires labialisées
non voisées fortes	*p	*t	*k	*q	*q <sup>w</sup>
éjectives		*t'	*k'	*q'	*q' <sup>w</sup>
Aspirées douces	*p <sup>h</sup>	*t <sup>h</sup>	*k <sup>h</sup>	*q <sup>h</sup>	*q <sup>wh</sup>

Après ce stade supposé, les éjectives, qui sont des consonnes postglottalisées non voisées, ont pu s'affaiblir en préglottalisées douces et voisées, avant de passer ensuite à des voisées simples non glottalisées<sup>40</sup>, comme dans la théorie standard.

Cette hypothèse glottalique présente plusieurs avantages. Tout d'abord son degré de plausibilité sur le plan typologique est nettement supérieur à celui de la

<sup>38</sup> De tels systèmes ont certes été observés, mais ils sont en tout cas rarissimes.

<sup>39</sup> Parmi les principaux promoteurs de l'hypothèse glottalique, on peut mentionner André Martinet, Morris Swadesh, Tamaz V. Gamkrelidze et Vyacheslav Ivanov, ainsi que Paul J. Hopper. Voir par exemple Paul J. Hopper, « *Glottalized and murmured occlusives in Indo-European* », *Glotta* 7, 1973, p. 141-166, André Martinet, *Des steppes aux océans*, op. cit., p. 161-166, et Tamaz V. Gamkrelidze et Vyacheslav Ivanov, *Indo-European and the Indo-Europeans*. Berlin : Mouton de Gruyter, 1995.

<sup>40</sup> André Martinet, op. cit., p. 164.

théorie standard ; en second lieu, elle rend compte de la rareté de \*b, car, comme l'écrivait fort justement André Martinet<sup>41</sup>, « l'inexistence d'une glottalisée dans l'ordre labial est un trait très fréquent dans les langues qui ont une série de ce type » (ce qui s'explique par des causes physiologiques). Il est donc normal que \*b soit rare si les consonnes voisées sont issues de glottalisées originelles, puisque l'on peut supposer qu'il n'y avait pas d'éjective labiale \*p' dans le proto-système (voir la case vide dans le tableau *supra*), donc pas de source initiale de \*b. Enfin, il est à noter que l'hypothèse glottalique rend mieux compte d'une contrainte imposant aux racines des mots indo-européens de ne pas à la fois commencer et se terminer par une plosive voisée simple<sup>42</sup> dans le cadre standard : cette contrainte se comprend bien mieux si l'on suppose qu'antérieurement il s'agissait d'empêcher en réalité la cooccurrence de deux consonnes glottalisées, propriété que l'on retrouve justement dans des langues modernes comprenant des éjectives dans leur système<sup>43</sup>.

Il serait facile de montrer comment, pour ce qui est de la phonation, le système postulé ci-dessus évoluera ensuite dans les différentes branches de la famille indo-européenne. Disons simplement que dans la majorité des principales branches, les développements des plosives ont été les suivants<sup>44</sup> :

- les non voisées fortes se sont maintenues telles quelles, sauf en germanique où elles se sont affaiblies en fricatives, et en arménien où elles se sont renforcées en aspirées (non voisées) ;
- les éjectives se sont transformées en voisées simples partout, sauf en germanique, en tokharien et en arménien, où elles donnent des non voisées simples ;
- les aspirées douces ont été préservées en sanskrit, se sont renforcées en aspirées fortes non voisées en grec, se sont à la fois renforcées et simplifiées en non voisées non aspirées en tokharien, simplifiées en voisées simples en hittite, avestique, germanique, arménien, celtique et balto-slave, et enfin ont abouti en latin à des résultats divers (avec perte générale de l'aspiration) dépendant largement du contexte ;
- le sanskrit a ajouté aux séries non voisée (p, t, etc.), voisée simple (b, d, etc.) issue des éjectives, voisée aspirée (b<sup>h</sup>, d<sup>h</sup>, etc.), une série voisée

---

<sup>41</sup> André Martinet, *op. cit.*, p. 163.

<sup>42</sup> Calvert Watkins, « *Proto-Indo-European: Comparison and Reconstruction* », in Anna Giacalone Ramat and Paolo Ramat (dir.), *The Indo-European Languages*. London & New York: Routledge, 1998, p. 53.

<sup>43</sup> Benjamin W. Fortson IV, *op. cit.*, p. 59.

<sup>44</sup> Pour un tableau général, voir par exemple Calvert Watkins, *op. cit.*, p. 35-37.

non aspirée (p<sup>h</sup>, t<sup>h</sup>, etc.) résultant de la fusion de séquences {plosive + h}<sup>45</sup>, ce qui lui vaut finalement quatre séries.

Dans tous les cas, ces développements s'expliquent aisément par les propriétés phonatoires au stade proto-indo-européen et par les pressions paradigmatiques ou le rôle du contexte, mais nous ne pouvons nous étendre plus longuement sur ce sujet dans les limites de cette contribution.

L'une des principales leçons à retenir des discussions – dont ce qui précède donne simplement une idée – est celle-ci : la connaissance des principes typologiques et des processus phonétiques naturels à l'œuvre dans les langues connues et attestées permet de mieux envisager la reconstruction d'une protolangue non attestée.

Quoi qu'il en soit, le lecteur aura compris que l'accord est une gageure dans l'entreprise de reconstruction. À cet égard, il convient de bien distinguer l'hypothèse d'une source commune, en l'occurrence d'une protolangue indo-européenne, et le résultat des procédures de reconstruction : la diversité des points de vue sur le système phonologique à reconstruire, par exemple, n'infirme en rien la validité générale de l'hypothèse selon laquelle on a affaire à une famille historique.

On ne constate pas davantage d'unanimité pour la flexion que pour la phonologie. Je me contenterai ici d'indiquer ce qui est le plus souvent admis dans le domaine nominal et le domaine verbal<sup>46</sup>. Pour le nom, on reconstruit normalement deux genres, l'animé et l'inanimé, avec une subdivision en masculin et féminin pour le genre animé ; le nom connaissait trois nombres : le singulier, le duel et le pluriel ; pour ce qui est du système casuel, on postule en général les mêmes cas qu'en sanskrit : nominatif, vocatif, accusatif, génitif, datif, ablatif, instrumental, locatif<sup>47</sup>.

En ce qui concerne le genre, on sait que le neutre correspond à l'inanimé, et que les formes de nominatif et d'accusatif du neutre connaissaient le syncrétisme, l'absence de forme distincte pour le nominatif s'expliquant par l'incapacité de principe d'avoir un sujet agent inanimé. Mais il est à noter de ce point de vue que le hittite avait développé un cas supplémentaire, l'ergatif, qui était

<sup>45</sup> André Martinet, *op. cit.*, p. 165.

<sup>46</sup> Benjamin W. Fortson IV, *op. cit.*, p. 75-139, et James Clackson, *op. cit.*, p. 90-156.

<sup>47</sup> La plupart des indo-européanistes présupposent que le sanskrit nous donne le système casuel maximal, mais rien ne garantit que le système casuel du sanskrit n'est pas lui-même le résultat de la fusion de cas anciennement distincts, ce qui donnerait un système proto-indo-européen encore plus complexe (James Clackson, *op. cit.*, p. 91-92). Inversement, d'autres penchent pour un système proto-indo-européen plus simple, qui se serait complexifié par la suite (Winfred P. Lehmann, *op. cit.*, p. 153-155).

affixé spécifiquement au nom neutre qui se trouvait en fonction de sujet d'une structure transitive. Par ailleurs, le hittite possédait aussi un allatif (appelé aussi directif) pour le lieu où l'on va, dévolu à l'accusatif dans les autres langues<sup>48</sup>.

Le verbe possédait quatre modes personnels : indicatif, impératif, subjonctif, optatif, deux modes non personnels : infinitif, participe, et trois bases flexionnelles pour le temps : présent, aoriste, parfait (l'imparfait étant un temps secondaire construit sur la base du présent) ; enfin, on postule deux voix : l'actif et le moyen (ou médiopassif).

Alors que les travaux d'Émile Benveniste sur la morphologie des racines indo-européennes font autorité<sup>49</sup>, la syntaxe<sup>50</sup> est sans doute le maillon le plus faible des tentatives de reconstruction, en raison même de la nature de cette composante du langage. En effet, l'opération est beaucoup plus problématique pour la phrase, qui est le produit de la créativité du locuteur, et on peut même s'interroger sur la faisabilité de la reconstruction syntaxique. On peut au mieux tenter de reconstruire les règles syntaxiques de production des phrases, mais les données présentent à cet égard beaucoup moins de garanties : tout d'abord, pour chaque point étudié, on peut en général postuler plusieurs hypothèses concurrentes sans qu'il soit possible de déterminer des critères sûrs permettant de les départager ; ensuite, au fil du temps, les différences entre les langues d'un même groupe s'accroissent bien davantage en syntaxe.

L'ordre des mots est l'un des sujets les plus étudiés : a-t-on affaire à un ordre SOV<sup>51</sup> en proto-indo-européen, par exemple ? S'agissait-il d'un ordre fixe, ou simplement de l'ordre le plus attendu, *ceteris paribus*, dans une langue où l'ordre des mots pouvait être relativement libre, comme en latin classique par exemple ?

Autre sujet très étudié : quels étaient les cas liés aux fonctions centrales sujet et objet direct ? Contrairement à ce que l'on pourrait penser, rien ne va de soi en cette matière : il suffit de songer à l'ergatif mentionné plus haut pour le hittite, qui démontre que le cas du sujet des propositions à mode personnel n'était pas forcément toujours le nominatif. Ajoutons qu'il est pratiquement impossible de savoir si l'ergatif du hittite représentait un état ancien de l'indo-européen, que les autres langues auraient perdu, ou s'il s'agissait d'une innovation propre au domaine anatolien en raison d'un contact de langues. Ce genre d'interrogation est

---

<sup>48</sup> Sur l'allatif et l'ergatif en hittite, voir Benjamin W. Fortson IV, *op. cit.*, p. 172.

<sup>49</sup> Émile Benveniste, *Origine de la formation des noms en indo-européen*, sixième tirage. Paris : Librairie d'Amérique et d'Orient, Jean Maisonneuve successeur, 1984.

<sup>50</sup> Pour d'excellentes considérations méthodologiques sur ce sujet, voir James Clackson, *op. cit.*, p. 157-186.

<sup>51</sup> S = sujet, O = objet direct, V = verbe.

caractéristique des tentatives de reconstruction syntaxique du proto-indo-européen.

Quittons à présent la question strictement linguistique pour passer à celle de la localisation du foyer indo-européen originel. L'hypothèse la plus en faveur est celle de la steppe pontique, qui s'étend d'ouest en est à partir du nord de la mer Noire, jusqu'au Kazakhstan occidental, en passant par la Russie méridionale. Cette hypothèse a été mise en valeur par les recherches de l'archéologue Marija Gimbutas (1921-1994), qui a défendu l'idée que les populations associées à la « culture des kourganes »<sup>52</sup> – terme russe désignant des tumulus – étaient en réalité les Indo-européens eux-mêmes. La théorie de Marija Gimbutas a été critiquée, notamment en raison de la diversité culturelle des peuples qu'elle a réunis sous cette appellation de « culture des kourganes », mais l'idée fondamentale d'incursions vers l'ouest, vers l'Europe, à partir des steppes, demeure celle qui est la plus largement retenue<sup>53</sup>.

La principale hypothèse concurrente est celle qui a été avancée par l'archéologue Colin Renfrew, qui situe le foyer en Anatolie<sup>54</sup> : elle n'a guère obtenu la faveur des indo-européanistes « orthodoxes » en raison notamment de problèmes de chronologie, car si l'on suit Renfrew, la dispersion du peuple primitif aurait eu lieu au moins trois millénaires trop tôt d'après les études mettant en relation la reconstruction lexicale et la technologie de la roue dont ce peuple a dû disposer pour ses migrations.

Quant aux autres hypothèses sur la localisation, elles ne sont guère crédibles. À ce propos, il est remarquable que certains persistent à défendre une position « indocentriste » en contradiction totale avec les données linguistiques. Selon la version faible de cette thèse, les Indo-aryens auraient toujours été basés en Inde, et selon la version forte, le berceau de l'indo-européen se situerait en Inde, les partisans d'une version encore plus radicale allant même jusqu'à identifier le sanskrit au proto-indo-européen, ou pire encore à nier toute différence entre dravidien et indo-aryen<sup>55</sup>.

---

<sup>52</sup> Marija Gimbutas, *The Kurgan culture and the Indo-Europeanization of Europe. Selected articles from 1952 to 1993*, *Journal of Indo-European Studies Monograph Series N° 18*, 1997.

<sup>53</sup> Benjamin W. Fortson IV, *op. cit.*, p. 45-46.

<sup>54</sup> Andrew Colin Renfrew, *Archaeology and Language: The Puzzle of Indo-European Origins*. London: Pimlico, 1987.

<sup>55</sup> Satya Swarup Misra, « The date of the Rigveda and the Aryan migration: fresh linguistic evidence », in Edwin F. Bryant et Laurie L. Patton (dir.), *The Indo-Aryan Controversy, Evidence and inference in Indian history*. London & New York: Routledge, 2005, p. 181-234. On relève par exemple l'assertion suivante, p. 208 : « Sanskrit is very close to Proto-Indo-European, if not identical ». Cet auteur affirme aussi, p. 226 : « The Dravidians are also Aryans ».



On ne peut pas présenter l'hypothèse indo-européenne sans mentionner la position critique de l'archéologue français Jean-Paul Demoule, qui en est l'un des adversaires les plus acharnés. Il expose son point de vue de manière détaillée dans son livre *Mais où sont passés les indo-européens ?*<sup>56</sup>. Selon l'auteur, les recherches indo-européennes sont à comprendre comme « la construction d'un mythe d'origine » destiné à concurrencer celui de la Bible<sup>57</sup>. Il mêle constamment le plan scientifique et les questions idéologiques ou politiques, s'attardant longuement sur « l'exploitation passée ou actuelle, par des mouvements nationalistes et extrémistes, du modèle indo-européen canonique [...] »<sup>58</sup> ; or, s'il est vrai que de telles exploitations ont effectivement eu lieu et qu'elles sont bien entendu condamnables et insupportables, elles ne sauraient pour autant décrédibiliser l'hypothèse indo-européenne sur le plan strictement scientifique.

Concernant la localisation du « peuple originel », Jean-Paul Demoule passe en revue les différentes possibilités en soulignant les écueils et les incohérences, mais il semble présupposer – à tort – que la difficulté que présente l'entreprise est inversement proportionnelle à la probabilité d'existence de ce peuple. Il met en avant la faiblesse des éléments de type archéologique, mais on pourrait lui objecter que la comparaison des langues anciennes attestées constitue une forme d'archéologie, qui nous assure peut-être des fondations plus solides.

Opposé à l'hypothèse d'une protolange, Demoule avoue sa préférence pour la position de Troubetzkoy, qui en 1928 proposait de distinguer les familles de langues (*Sprachfamilien*) et les affinités linguistiques aréales (*Sprachbünde*)<sup>59</sup>. L'un des exemples les plus connus et les plus convaincants d'affinités aréales est celui de la région comprenant l'albanais, le grec, le roumain et le bulgare, et dans une moindre mesure le serbe<sup>60</sup>, qui appartiennent à quatre groupes historiques distincts de l'indo-européen, mais présentent néanmoins certaines similarités morphosyntaxiques possiblement dues à la proximité géographique. Plus tard, dans ses « Réflexions sur le problème indo-européen »<sup>61</sup>, Troubetzkoy a appliqué cette notion d'affinité aréale à l'indo-européen, affirmant que les langues en question

---

<sup>56</sup> Jean-Paul Demoule, *Mais où sont passés les indo-européens ? Le mythe d'origine de l'Occident*. Paris : Seuil, 2014.

<sup>57</sup> *Op. cit.*, p. 596.

<sup>58</sup> *Op. cit.*, p. 598.

<sup>59</sup> *Op. cit.*, p. 163-168.

<sup>60</sup> Le bulgare et le serbe sont des langues slaves, le roumain est une langue romane, et quant au grec et à l'albanais, ils appartiennent chacun aux groupes éponymes.

<sup>61</sup> Nicolas S. Troubetzkoy, « *Gedanken über das Indogermanenproblem* », *Acta Linguistica*, volume 1, issue 1, 1939 ; *Correspondance avec Roman Jakobson et autres écrits*, édition établie par Patrick Sériot, traduit du russe par P. Sériot & M. Schönerberger. Payot : Lausanne, 2006.

résultent non pas de la fragmentation d'une protolangue, mais de l'évolution convergente de langues de familles historiques distinctes. Or, s'il est vrai que telle ou telle correspondance phonologique peut ne s'appliquer qu'à un sous-ensemble des langues concernées et que le lexique fondamental est rarement commun au domaine tout entier, cela ne suffit pas à nier la systématique qui caractérise globalement les relations observées, et qui discrédite la position de Troubetzkoy : cet immense phonologue a manifestement commis là une sérieuse erreur.

Nous en arrivons à la dernière partie de cette contribution. Certains comparatistes ont proposé des hypothèses encore plus hardies que la théorie standard ou ses variantes concurrentes, soit en postulant des relations entre la protolangue indo-européenne et les protolangues d'autres familles, comme la famille ouralique, par exemple, ou encore le sémitique et le dravidien, soit en postulant une superfamille dont l'indo-européen ne serait qu'une branche. L'hypothèse nostratique<sup>62</sup> est la première entreprise du genre. Le nostratique serait une vaste famille incluant l'indo-européen, l'ouralique, le sémitique, le dravidien et le kartvélien (géorgien et langues apparentées du Caucase). L'idée vient du linguiste danois Høger Pedersen (1867-1953) au début du XX<sup>e</sup> siècle et a été reprise dans les années 1960 par les linguistes soviétiques Vladislav Illich-Svitych (1934-1966) et Aharon Dolgopolsky (1930-2012).

Une théorie très proche est celle de la superfamille eurasiatique, qui est ancienne mais a été défendue ces dernières décennies par le linguiste américain Joseph Greenberg (1915-2001)<sup>63</sup>. Cette superfamille serait encore plus vaste que le nostratique.

La plupart des indo-européanistes sont loin d'être convaincus par ces approches et la méthode qui les sous-tend, préférant s'en tenir à la théorie standard ou à une approche équivalente<sup>64</sup>.

Je voudrais conclure par une réflexion sur la nature des données anciennes attestées et étudiées par les indo-européanistes : ce sont des traces écrites, mais

---

<sup>62</sup> De l'adjectif latin *nostrās*, « qui est de chez nous », dérivé de *noster*, « notre ».

<sup>63</sup> Joseph H. Greenberg, *Indo-European and Its Closest Relatives: The Eurasiatic Language Family*. Volume 1, Grammar. Stanford: Stanford University Press, 2000, et *Indo-European and Its Closest Relatives: The Eurasiatic Language Family*, Volume 2, Lexicon. Stanford: Stanford University Press, 2002. Merritt Ruhlen est le disciple de Greenberg qui a poussé le réductionnisme en matière de familles de langues jusqu'à ses limites extrêmes : voir Merritt Ruhlen, *The Origin of Language: Tracing the Evolution of the Mother Tongue*. New York: John Wiley and Sons, 1994, et *On the Origin of Languages: Studies in Linguistic Taxonomy*. Stanford: Stanford University Press, 1994.

<sup>64</sup> Benjamin W. Fortson IV, *op. cit.*, p. 13.

qu'est-ce qu'une trace écrite ? C'est l'inscription de signes censés figurer des sons, lesquels sont à jamais perdus. Autrement dit, on travaille indirectement sur des sons que l'on n'a jamais entendus, et l'on essaie d'en reconstituer d'autres plus anciens encore – leurs ancêtres en quelque sorte –, jamais entendus non plus par nous. Ces sons, tout comme le lexique et la grammaire, s'organisent en systèmes intériorisés mentalement par les locuteurs<sup>65</sup> et transformés graduellement et imperceptiblement de génération en génération, et nous sommes nous-mêmes des maillons de cette chaîne. Quand on envisage ainsi l'histoire du langage, il apparaît plus clairement que les linguistes qui ont relevé le défi en tentant de retisser la trame de l'indo-européen à partir de données aussi fragiles méritent toute notre admiration.

---

<sup>65</sup> Ce sont ces systèmes intériorisés mentalement par les locuteurs, et en réalité inobservables en tant que tels, que l'on appelle *langues* : ce qui est observable, c'est la partie audible ou visible du discours, mais en aucun cas la langue.