

# La reconstruction de l'indo-européen et la réalité du sémitique : convergences et perspectives

**Arnaud Fournet**

## La création de l'indo-européen

Le terme *indo-européen* est attesté pour la première fois sous la plume de Thomas Young en 1813. Mais la parenté indo-européenne a été reconnue beaucoup plus tôt, dès le XVI<sup>e</sup> siècle. La ressemblance entre sanscrit, latin et grec fut notée pour la première fois en 1583 par un jésuite anglais, Thomas Stephen, qui vécut en Inde de 1579 à 1619. Des gens ayant des intérêts plus terrestres, comme le marchand italien Filippo Sassetti, en 1585, furent frappés par la familiarité du sanscrit et sa similitude avec le latin et le grec. Sassetti cite en particulier les nombres de 6 à 10, ainsi que les mots « Dieu » et « serpent ». Néanmoins, il n'a jamais formulé l'hypothèse d'une parenté, car il ressort de son courrier une sorte d'effarement devant l'écart culturel qui sépare l'Italie de l'Inde. Beaucoup de travail fut accompli, spécialement aux Pays-Bas par Marcus Boxhorn (1640) et en France par Claude de Saumaise (1643), sur le lexique des langues indo-européennes, qui n'avaient pas alors ce nom, surtout sur le sanscrit, le grec, le latin, le perse et les langues germaniques. La démarche était alors strictement comparative et, dans cette époque très fixiste, la notion de proto-langue était tout simplement impensable. En outre, en l'absence de données précises sur le sanscrit et les langues perses, la démonstration d'une parenté restait peu solide, voire presque fabuleuse.

Les similitudes évidentes de ces langues furent expliquées dans le cadre de l'origine « scythique », parfois aussi appelée « japhétique ». Le peuple scythe bien connu, une branche iranienne de l'indo-européen, fut alors supposé avoir essaimé à travers toute l'Eurasie et s'être ramifié en autant de langues modernes. Leibniz (1646-1716) a contribué à propager cette hypothèse d'une diffusion « scythique » qui remonte en fin de compte à Boxhorn<sup>1</sup> :

On peut conjecturer que cela vient de l'origine commune de tous ces peuples descendus des Scythes, venus de la mer Noire, qui ont passé le Danube et la Vistule, dont une partie pourrait être allée en Grèce, et l'autre aura rempli la Germanie et les Gaules.

D'autres travaux, comme ceux de James Parsons (1767), sont intéressants sur le plan comparatif, spécialement en ce qui concerne la ressemblance entre langues celtes – irlandais et gallois –, mais ils sont entrelacés de façon inextricable avec une thématique biblique, relative à la Genèse, de sorte qu'on peine à distinguer

---

1. Leibniz (1990, p. 218).

les hypothèses véritablement historiques et les éléments plus mythiques, voire fantastiques<sup>2</sup>.

Les mots *japhétique* et *indo-européen* sont restés longtemps en concurrence. En 1905, on écrit encore :

Le sujet du présent livre est le groupe de langues que les Allemands appellent aujourd'hui indo-germanique (idg.), et que l'on désigne aussi sous le nom d'indo-européen (i.e.), nom usuel en France et qui sera adopté dans la présente traduction, ou d'aryen, ou de japhétique<sup>3</sup>.

Avant l'invention des notions d'évolution et de préhistoire<sup>4</sup>, les différences furent expliquées non pas par des divergences progressives au fil du temps, mais par des mélanges entre langues dans des proportions variables. Dans ce cadre prémoderne, les langues n'évoluent pas mais elles se mélangent, ce qui donne naissance à d'autres idiomes. Un point de vue clairement moderne sur l'évolution des langues est exprimé par Jakob Grimm (1785-1865) dans *Geschichte der Deutschen Sprache* (1848, p. 833) :

Tous les dialectes se développent dans un ordre progressif, et plus on remonte vers l'origine des langues, plus leur nombre diminue et plus leurs différences s'effacent. S'il n'en était pas ainsi, la formation des dialectes et la pluralité des langues resteraient inexplicables. Toute diversité est sortie graduellement d'une unité primitive. Les dialectes allemands se rapportent tous à une ancienne langue germanique commune, et celle-ci à son tour, à côté du lithuanien, du slave, du grec et du latin, n'était qu'un des dialectes d'un idiome primitif plus ancien encore.

Une autre présentation moderne de la parenté indo-européenne est formulée par Friedrich Schlegel dans *Ueber die Sprache und die Weisheit der Indier* (1808) :

Le sanscrit présente un lien de parenté très fort avec le latin, le grec et le germanique ainsi que le perse. Les similitudes ne se limitent pas seulement à un très grand nombre de racines que ces langues ont en commun mais existent aussi dans la structure et la grammaire. En conséquence de quoi le rapprochement n'est pas accidentel, explicable par des échanges, mais fondamental, provoqué par une origine commune.

Le mérite de la démonstration de la parenté revient à Franz Bopp (1791-1867). Venu étudier le sanscrit à Paris en 1812, il publie quatre ans plus tard un ouvrage d'une ampleur colossale, qui formalise de façon définitive l'apparement qui est dans les têtes érudites depuis 350 ans. En relisant la *Grammaire comparée des langues indo-européennes*, on reste admiratif devant le travail accompli. En fait, presque tout ce qui fait le comparatisme y est déjà !

---

2. Pour en savoir plus, on pourra consulter les ouvrages d'érudition scientifique suivants : Sergent (1995, p. 20-46), synthèse faite par un historien ; Mallory (1997a), synthèse faite par un archéologue.

3. Brugmann (1905, p. 2). La traduction française de ce livre commence par cette phrase.

4. Ce mot est attesté en français seulement à partir du milieu du XIX<sup>e</sup> siècle.

Un mot doit être dit à propos de Sir William Jones (1746-1794), qui est souvent présenté comme le découvreur du sanscrit et comme l'initiateur du comparatisme indo-européen. En 1786, Jones, qui était juge à la cour suprême de Calcutta, prononça un discours auprès de la Société royale asiatique du Bengale, dont nous traduisons un extrait demeuré célèbre :

Le sanscrit, quelle que soit son antiquité, est d'une structure merveilleuse, plus parfaite que le grec, plus riche que le latin et plus subtilement raffinée que ces deux derniers, tout en ayant avec eux une affinité si forte dans les racines des mots et dans les formes grammaticales qu'elle ne saurait s'être produite par hasard, si forte en effet qu'aucun philologue ne pourrait les examiner toutes les trois sans croire qu'elles ne viennent de quelque origine commune, qui *peut-être* n'existe plus. Il existe une raison similaire, bien que moins contraignante, pour supposer que le gotique et le celte, quoique *mélangés* à un idiome différent, ont la même origine que le sanscrit, et le vieux perse pourrait être ajouté à la famille.

Nous avons souligné deux expressions qui montrent les limites de la modernité de Jones. Il reste dans le cadre ancien du mélange de langues et ne franchit pas le cap d'une origine commune si ancienne qu'elle serait perdue et à reconstruire. Jones est un érudit, certes bien informé mais qui représente la fin d'une époque, mélangiste et fixiste. Son apport réel au comparatisme indo-européen est très surestimé. Ses compétences linguistiques également, car il a considéré que le farsi avait la même origine que l'arabe, puisqu'écrits dans le même alphabet. Il est instructif de citer la phrase qui précède immédiatement le paragraphe demeuré célèbre :

Le pur hindi, qu'il soit d'origine tartare ou chaldéenne, était la langue originale de l'Inde du Nord, dans laquelle le sanscrit fut introduit par des conquérants venus d'autres royaumes à quelque époque reculée, car nous ne pouvons pas douter que la langue des Védas a été utilisée dans la majeure partie du pays, aussi longtemps que la religion des Brahmanes l'a dominé.

En clair, Jones ne voit pas la filiation historique entre le hindi et le sanscrit. Pour lui, le hindi serait d'origine altaïque ou sémitique et antérieur au sanscrit. Il va de soi que les thuriféraires anglophones de Jones ne mentionnent guère ces assertions suspectes de leur champion. En général, on ne cite que le paragraphe le plus favorable, soigneusement extrait de son contexte et ébavuré.

En outre, faire démarrer aussi tard – à la fin du XVIII<sup>e</sup> siècle – les recherches sur l'indo-européen aboutirait à en faire une des familles les plus tardivement identifiées, un demi-siècle après la famille ouralienne par exemple<sup>5</sup>. C'est tout simplement absurde.

---

5. La première affirmation claire d'une parenté ouralienne est due à Philip Johann van Strahlenberg, un officier suédois, en 1729, qui s'est exprimé en langue allemande. La première traduction en français, publiée à Amsterdam, est de 1757 : *Description de l'empire russe*. Voir le tome 1, p. 148 et suivantes.

Les sous-familles reconnues pour l'indo-européen sont les suivantes :

- celtique et italique,
- germanique,
- albanais,
- thrace, dace (éteintes),
- grec, arménien, phrygien (éteint),
- anatolien : hittite, louvite, lydien (connues depuis 1915),
- balte et slave,
- indo-iranien,
- tokharien.

### **Le formalisme indo-européen**

Quand on part d'un corpus de données lexicales, la première étape de travail consiste à trier et à comparer afin de mettre au jour les faits intéressants. En premier lieu, il faut repérer un certain nombre de cognats potentiels et poser un certain nombre de correspondances. Cette phase d'analyse débouche normalement sur une nouvelle phase appelée « reconstruction ». Ces travaux ont commencé pour l'indo-européen vers le XVI<sup>e</sup> siècle et ils ont gagné en rigueur et en précision au début du XIX<sup>e</sup> siècle.

On reconstruit des étymons ou des proto-formes qui possèdent plusieurs caractéristiques :

- ils sont censés remonter à la proto-langue, qui existait autrefois,
- ils synthétisent les résultats de l'analyse comparative,
- ils sont formatés de manière à pouvoir en dériver les formes attestées.

Dans les présentations usuelles qui sont faites de la méthode comparative, on parle très peu du *format* dans lequel doit se mouler la proto-langue. Cette notion est à notre avis essentielle. Le format utilisé pour la reconstruction de l'indo-européen a évolué plusieurs fois depuis le début du XIX<sup>e</sup> siècle. Ainsi le mot *père* a été reconstruit :

- \**pitâr* quand on considérait le sanscrit comme étant la proto-langue,
- \**pâter* dans un formalisme empirique,
- \**p<sub>2</sub>ter* dans un formalisme semi-structuraliste,
- \**pH<sub>2</sub>ter* dans un formalisme structuraliste et laryngaliste.

Avec ces différents formats, on rend compte des mêmes données primaires, à savoir : latin *pater*, grec *patêr*, sanscrit *pitâr*, irlandais *aithir*, etc. Notre avis est qu'il est important de comprendre que le formalisme utilisé pour rendre compte des données n'est pas dans les données et qu'il n'est pas neutre vis-à-vis de la conception que l'on se fait de la proto-langue et du cadre théorique dans lequel se placent les reconstituteurs. Certains esprits narquois ont souligné que le visage de l'indo-européen évoluait beaucoup plus vite que celui des langues-filles. C'est directement lié aux progrès du comparatisme et de la linguistique elle-même. Le format de la reconstruction évolue parce qu'on pose sur la proto-langue un regard qui évolue. Ces différents formats ne sont pas plus vrais les uns que les au-

tres, ni mutuellement exclusifs. Ils ne disent pas la même chose. Bien sûr, on attend du format le plus récent qu'il dise plus et mieux que les précédents.

Le premier âge de la reconstruction \**pitar*, avant 1840, ne se projetait pas encore dans les notions de proto-langue et de préhistoire. Il avait le sanscrit comme horizon indépassable. Le deuxième âge \**pāter* est un premier pas vers la notion de langue préhistorique. Il reconstruit des étymons mais il reste très près des données, ce qui donne une proto-langue avec un bric-à-brac de racines présentant des alternances morphologiques très variées : e ~ o ~ Ø (zéro), ê ~ ô ~ ə, â ~ ô ~ a, etc. Le formalisme de cette époque est en outre très *vocalique*. Ensuite, les deux formats suivants sont des perfectionnements significatifs résultant d'une réflexion en profondeur sur la phonologie et la structure des racines de l'indo-européen. Ce sont les points que nous allons maintenant examiner de plus près.

De façon générale, l'indo-européen est tiraillé entre deux pôles : les comparatistes philologues se contentent d'un mécano phonético-algébrique alors que les comparatistes linguistes veulent une proto-langue réaliste, acceptable. D'un côté, on veut des formules et des règles de réécriture ; de l'autre, on veut un état de langue possible. Les évolutions du formalisme indo-européen se font pour concilier les deux points de vue et pour répondre à deux objectifs principaux : améliorer l'efficacité du mécano algébrique et renforcer le réalisme de la proto-langue. Nous allons voir que toutes les réformes à ce jour aboutissent à rapprocher l'indo-européen du sémitique.

Le système des consonnes postulé au tout début était calqué sur le sanscrit. La structure parfaite et rectangulaire des occlusives du sanscrit avait été soulignée dès l'Antiquité par les Indiens eux-mêmes, et en particulier par Panini, et elle avait séduit les premiers comparatistes :

	Labiales	Dentales	Rétroflexes	Palatales	Vélaires
Sourdes	p	t	T	č	k
Sourdes aspirées	p <sup>h</sup>	t <sup>h</sup>	T <sup>h</sup>	č <sup>h</sup>	k <sup>h</sup>
Sonores	b	d	D	j	g
Sonores aspirées	b <sup>h</sup>	d <sup>h</sup>	D <sup>h</sup>	j <sup>h</sup>	g <sup>h</sup>
Nasales	m	n	N	ń	

Ce tableau a servi de crible au XIX<sup>e</sup> siècle pour trier les cognats légitimes et le sanscrit a servi de fil directeur pour mettre en place le mécano algébrique des reconstructions. Il suffisait de voir à quoi correspondaient les sons du sanscrit dans les autres langues indo-européennes, et le tour était joué.

Il est vite apparu que la série des sourdes aspirées était marginale sur le plan quantitatif et qu'elle concernait surtout le sanscrit lui-même. Le système orthodoxe, une fois supprimée la série sourde aspirée, repose sur trois séries fondamentales : sourde, sonore, sonore aspirée. Du point de vue du mécano algébrique des comparatistes, cela ne change rien. Mais du point de vue linguistique, c'est gênant. Car ce système de traits n'existe nulle part sur la planète et il n'est

pas acceptable par la phonologie. A partir de 1950, cela a provoqué la recherche d'un nouveau format plus acceptable pour le proto-système phonologique : les systèmes dits « glottaliques ».

Système orthodoxe		Système glottalique	
<i>Etait sourde</i>	t	<i>Devient sourde</i>	t
<i>Etait sonore</i>	d	<i>Devient glottalisée</i>	t̚ [tʔ]
<i>Etait sonore aspirée</i>	d <sup>h</sup>	<i>Devient sonore</i>	d

Le reformatage des traits ne change pas le nombre de phonèmes ni les correspondances. Il propose de changer la nature des traits du proto-système. Il existe plusieurs variantes de système glottalique<sup>6</sup> mais, par rapport à notre propos, ils aboutissent en pratique à aligner l'indo-européen sur le système de l'hébreu et du proto-sémitique : à savoir *sourde*, *sonore*, *sourde glottalisée* (ou *sourde emphatique*). Ce système est plus acceptable puisqu'on est certain qu'il existe !

Un autre argument en faveur des théories glottaliques est la rareté troublante de \*b, pour lequel on ne reconstruit que très peu de lexèmes. On sait que le phonème /pʔ/ n'existe pas dans tous les systèmes à glottalisation, car la coarticulation labiale et glottale est peu efficace. La réinterprétation de \*b comme valant en fait \*/pʔ/ expliquerait donc sa rareté.

A ce jour, ce visage glottalique de l'indo-européen n'est pas accepté de façon universelle comme valide et fait partie des hypothèses qui clivent les comparatistes en deux camps. Celui des sceptiques reste majoritaire à ce jour. Par exemple, Michael Meier-Brügger, qu'on peut considérer comme représentatif du comparatisme orthodoxe de langue allemande, affirme<sup>7</sup> : « *Die Zweifel an der Berechtigung des Ansatzes von Glottalen für das Uridg sind nach Ansicht der Verfasser des Studienbuches nicht ausgeräumt.* » Indépendamment des réticences théoriques, sur le plan pratique, si l'on mettait en œuvre le format glottalique dans la reconstruction de l'indo-européen, il faudrait réécrire 150 ans de travaux accumulés. Une entreprise titanesque...

D'autres auteurs adhèrent et vont même jusqu'à préciser que ces glottalisées seraient des pré-glottalisées, par exemple Robert Beekes<sup>8</sup> : « *Important is that the glottalic feature probably preceded the consonant.* »

Les glottalisées se manifestent en slave comme des préglottalisées. En effet, slave /d/ < \*[ʔt] est toujours précédée d'un /ê/ long, alors que slave /d/ < \*[d<sup>h</sup>] est toujours précédée d'un /e/ bref. L'explication en est que [\*v+\*ʔt] > [\*v+\*ʔd] >

6. Pour un panorama des hypothèses : Collinge (1985, Appendix II, p. 259-269) ou Salmons (1993).

7. Meier-Bruegger (2002, p. 126). Nous traduisons : « Les doutes sur la validité de l'emploi de glottal(isé)es pour l'indo-européen ne sont pas éliminés, selon l'avis de l'auteur du [présent] manuel. »

8. Beekes (1995, p. 133). Nous traduisons : « Il est important de noter que le trait glottalique précédait probablement la consonne. »

[\*v<sup>2</sup>+\*d], d'où une voyelle longue<sup>9</sup>. Ce phénomène porte le nom éponyme de *Loi de Winter*, d'après le découvreur. Certains comparatistes en concluent que tout le domaine indo-européen avait des préglottalisées. Nos propres travaux, dont il serait trop long de faire état ici, sur la séquence *arrêt glottal + sourde*, montrent que seules les langues centrales – germanique, balte, slave, grec, arménien, indo-iranien – ont des préglottalisées. Le celtique et l'italique ont des postglottalisées. À l'appui de la théorie glottalique, certains comparatistes, comme Kortlandt, soulignent que certaines langues indo-européennes présentent encore de nos jours des glottalisées, comme par exemple les variétés d'anglais parlées en Angleterre même. De notre point de vue, il est intéressant de confronter ces hypothèses glottaliques avec la réalité concrète du sémitique, qui est la terre d'élection des glottalisées-emphatisées. Nous donnerons deux exemples. La racine du verbe *donner* √ dō- contient un /d/ initial. Si l'on admet que /d/ vaut \*/ṭ/ [tʔ], à savoir ṭ ou v, il est flagrant que le verbe « donner », en arabe 'aṭā' et en hébreu *naṭan*, contient une emphatique (ou glottalisée) radicale, conforme à ce qu'on attend d'après le reformatage de l'indo-européen. Réciproquement, le mot « allaiter, têter » √ dheH<sub>1</sub> correspond à hébreu *dad*, « sein », sans glottalisée de part et d'autre<sup>10</sup>.

En résumé, pour faire simple, nous dirons que le reformatage glottalique de l'indo-européen est nécessaire pour des raisons internes au domaine indo-européen et qu'il offre en prime des perspectives macro-comparatives très prometteuses.

En ce qui concerne les voyelles, de toutes les alternances possibles dans un squelette de consonnes, il est vite apparu que le paradigme e ~ o ~ Ø était la référence absolue : le cas le plus fréquent et le plus régulier. Les comparatistes appellent *apophonie* l'alternance e ~ o ~ Ø. /e/ est appelé degré e, /o/ degré o et Ø degré zéro. Il existe dans les langues indo-européennes des exemples très clairs de cette apophonie e ~ o ~ Ø héritée :

- latin *ped-is* ~ grec *pod-os* « pied » au cas génitif ;
- latin *tegō* « je couvre » ;
- latin *tegumen* « tégument, peau » (< qui couvre la chair) ;
- latin *tect-um* « toit » (< qui couvre la maison) ;
- latin *toga* « toge » (< qui couvre le corps).

En latin et en grec, il y a cinq voyelles brèves : a, e, i, o, u. Mais elles n'ont pas le même statut : /i/ et /u/ ne sont pas intégrées dans un jeu d'alternances morphologiques. La voyelle /a/ existe dans toutes les langues mais elle ne rentre pas non plus dans des alternances normales et régulières. Comme Antoine Meillet l'a fait remarquer avec pertinence, /a/ est fréquent surtout à l'initiale et il alterne avec

9. Beekes (1995, p. 133). "The solution is that the glottal stop /ʔ/ lengthens the preceding vowel : a?g-t > âc-t. [...] The theory also explains why in Balto-Slavic a preceding vowel is lengthened by a voiced consonant (Winter's Law) [...] It is now possible to make out whether a Balto-Slavic \*g goes back to an aspirated [voiced] or a non-aspirated [voiced] sound."

10. Le rapprochement avec l'arabe *naḥd* est plus délicat (double incrémentation...), mais intéressant.

/o/ et jamais /e/. Nous en verrons plus loin les raisons avec la structure des racines<sup>11</sup>.

Dans les verbes, /e/ a valeur de présent et /o/ a valeur de passé. Ce système se voit bien en grec et en germanique, y compris dans une langue moderne comme l'anglais. L'alternance e ~ o est souvent reflétée sous une forme évoluée : i (< \*e) ~ a (< \*o), en germanique.

	e	o	∅
Valeur	Présent	Passé	Dérivé
Anglais <i>to get</i> « obtenir »	get	got	
Anglais <i>to bear</i> « porter »	bear [ber]	bore [bor]	
Anglais <i>to be born</i> « naître »		born [born]	birth [bərθ]
Anglais <i>to drink</i> « boire »	drink	drank	drunk
Anglais <i>to bring</i> « apporter »	bring	brought [bro:t]	

### L'invention des laryngales

L'alternance e ~ o ~ ∅ marche bien tant que la racine a un beau squelette de consonnes C<sub>1</sub>-C<sub>2</sub>. Elle est dérégulée dans certaines racines, comme par exemple celle du verbe *donner*. En latin, l'entrée du dictionnaire liste les formes : *dô* « je donne », *dâs* « tu donnes », *dâre* « donner », *dedi* « j'ai donné », *datum* « donné » et il faut ajouter *dônum* « don ». On voit mal comment intégrer ces formes dans le même paradigme que les autres verbes. C'est aussi le cas de la racine « être debout », qui a en grec un ensemble de formes *sta*, *stâ*, *stau*, *stû*, morphologiquement opaques. L'intégration de ces racines aux alternances étranges dans un modèle unifié est un autre problème de la reconstruction.

Dans son *Mémoire* sur les voyelles de l'indo-européen, écrit en 1870 à 18 ans (!), Ferdinand de Saussure a étudié de plus près les voyelles longues du grec ancien et a confirmé qu'il n'existait qu'une seule alternance morphologique fondamentale : e ~ o ~ ∅. Et il a postulé des *coefficients sonantiques* pour rendre compte des autres alternances et des voyelles longues. Dans l'approche de Saussure, il en faut trois, notées E, A, O :

Données	Format saussurien avec coefficients <sup>12</sup>
ê ~ ô ~ e	eE ~ oE ~ ∅E
â ~ â ~ a	eA ~ oA ~ ∅A
ô ~ ô ~ o	eO ~ oO ~ ∅O
stâ ~ sta	steA ~ stA
dô ~ do	deO ~ dO

11. #a- < \*H<sub>2</sub>e et #e- < \*H<sub>1</sub>e sont en distribution complémentaire.

12. Voir Saussure (1870, p. 145) pour un tableau similaire.



Si l'on raisonne en termes de traits ou d'effets phonétiques, A colore en voyelle /a/ et allonge, E allonge et ne colore pas, O allonge et colore en /o/. En suivant Saussure, la racine « donner » est reformatée d\_O et la racine « être debout » st\_A. L'approche de Saussure est plus structuraliste et plus abstraite. Elle régularise la morphologie mais elle pose le problème de déterminer quel genre de réalité phonétique se cache derrière ces *coefficients*, postulés un peu *ex nihilo* pour les besoins de la cause. Tant qu'on se contente d'un mécano graphique, ce genre de format fonctionne, mais si l'on veut donner un sens linguistique à ces *coefficients sonantiques*, la situation devient plus problématique.

Notons au passage que Saussure s'en est toujours tenu à une interprétation vocalique des coefficients, alors que Hermann Möller est incontestablement le premier à avoir postulé une nature consonantique pour ces coefficients<sup>13</sup>.

Par ailleurs, il est apparu que le degré dit zéro (Ø) avait souvent un reflet vocalique au lieu d'être amui et strictement égal à zéro. Le sanscrit présente alors /i/ là où les autres langues ont /a/. Sur le modèle de l'hébreu, qui présente des alternances *schwa* ~ *voyelle pleine*, les comparatistes ont appelé *schwas*, notés /ə/, les trois reflets vocaliques des coefficients sonantiques de Saussure. Du point de vue de l'épistémologie des sciences et de l'histoire du comparatisme, il est surprenant que personne n'ait jamais étudié l'importation du concept de *schwa hébreu* dans le domaine indo-européen. Il est pourtant manifeste que les premiers comparatistes connaissaient les langues majeures de l'Antiquité européenne : grec, latin et hébreu...

Dans ce format semi-vocalique de l'indo-européen, les alternances E A O sont configurées ainsi :

Données	Format saussurien avec coefficients	Format empirique avec schwas <sup>14</sup>
ê ~ ô ~ e	eE ~ oE ~ ØE	eə <sub>1</sub> ~ oə <sub>1</sub> ~ ə <sub>1</sub>
â ~ ô ~ a	eA ~ oA ~ ØA	eə <sub>2</sub> ~ oə <sub>2</sub> ~ ə <sub>2</sub>
ô ~ ô ~ o	eO ~ oO ~ ØO	eə <sub>3</sub> ~ oə <sub>3</sub> ~ ə <sub>3</sub>
stâ ~ sta	steA ~ stA	steə <sub>2</sub> ~ stə <sub>2</sub>
dô ~ do	deO ~ dO	deə <sub>3</sub> ~ də <sub>3</sub>

Le but ultime de la démonstration de Saussure dans le *Mémoire sur le système primitif des voyelles dans les langues indo-européennes* est résumé en page 135 :

Le phonème a<sub>1</sub> [moderne \*e] est la voyelle radicale de toutes les racines. Il peut être seul à former le vocalisme de la racine ou bien être suivi d'une seconde sonante que nous avons appelé coefficient sonantique (p. 8.) [...]

13. Szemerényi (BSL, n° 68, 1973, p. 7) : « En net contraste avec Saussure, Möller essaya, dès le début, de donner des définitions phonétiques de ces consonnes perdues. »

14. La numérotation 1 pour [e] et 2 pour [a] remonte directement à Brugmann, par l'entremise de Saussure.

Les phonèmes A et O [moderne \*H<sub>2</sub> et \*H<sub>3</sub>] sont des coefficients sonantiques. Ils ne pourront apparaître à nu que dans l'état réduit [zéro] de la racine. A l'état normal de la racine, il faut qu'ils soient précédés de a<sub>1</sub> [moderne \*e], et c'est des combinaisons a<sub>1</sub>+A, a<sub>1</sub>+O, [moderne \*e+H<sub>2</sub> et \*e+H<sub>3</sub>], que naissent les longues â et ô.

De même en page 141 :

L'ê long, dans notre théorie, ne doit pas être un phonème simple. Il faut qu'il se décompose en deux éléments. Lesquels ? Le premier ne peut être que a<sub>1</sub> (e). Le second, le coefficient sonantique, doit apparaître à nu dans la forme réduite (p. 135)<sup>15</sup>.

Du point de vue pratique, ces trois formats – vocalique, coefficient ou schwa – donnent les mêmes résultats, mais sur le plan linguistique, ils sont très différents. Tant qu'on ne s'interroge pas sur la nature phonétique possible de ces *entités*, on peut en rester à une stricte algèbre graphique, proche du jeu d'écriture<sup>16</sup> :

Die Bezeichnung Laryngale für den Ansatz dreier Konsonanten (Engelaute) der idg Grundsprache ist wissenschaftsgeschichtlich bedingt. Für diese ist eine algebraistische Notierung als uridg \*h<sub>1</sub>, \*h<sub>2</sub> bzw \*h<sub>3</sub> üblich geworden. [Notre traduction : L'appellation *laryngale* pour trois consonnes de l'indo-européen est scientifiquement certaine. Pour elles, une notation algébrique \*h<sub>1</sub>, \*h<sub>2</sub> et \*h<sub>3</sub> est devenue usuelle.]

La logique interne de la morphologie arabe amène à la conclusion que le [â] long repose sur une séquence implicite \*[awa] attestée nulle part, dans les synchronies arabophones. Ici \*/w/ est un *fantôme* morphologique, que la logique de la langue oblige à postuler pour rétablir la régularité là où la synchronie ne voit qu'anomalie. De la même façon, la logique interne de l'indo-européen oblige à postuler au moins trois entités phonologiques supplémentaires qui permettent de retrouver un point de fonctionnement stable et cohérent. La démarche intellectuelle de Saussure est *in fine* identique à celle des grammairiens arabes, tout en étant complètement indépendante et appliquée à un corpus de données totalement différent.

Outre leurs effets sur la longueur et le timbre vocalique, Saussure lui-même, dès 1870, avait noté que ces *fantômes* ont des effets sur les consonnes. Ainsi, la racine \*√ st\_A au degré zéro \*stA aboutit en sanscrit à √ st<sup>h</sup> au lieu de √ st. La racine « boire » \*√ p\_O au degré zéro \*pO aboutit en latin à une sonore /b/ dans *bibere*, d'où le français « boire ». Ces *coefficients-schwas* ont un impact sur les consonnes : ils provoquent aspiration ou sonorisation<sup>17</sup>. Ils ont plusieurs propriétés phonétiques observables : allongement, coloration, aspiration, sonorisation... On comprend qu'ils ont longtemps eu un parfum d'hérésie ou de scandale.

---

15. On notera que Saussure postule bien trois coefficients sonantiques pour la morphologie. Certains auteurs prétendent qu'il n'en aurait postulé que deux. Ils n'ont pas lu le *Mémoire*.

16. Meier-Bruegger (2002, p. 106, § L314).

17. Outre Saussure, le Danois Holger Pedersen (1867-1953) et Albert Cuny (1870-1947) sont arrivés à des conclusions identiques.

Les comparatistes qui ont adhéré dès 1870 à la notion de coefficients sonantiques ont été contraints de réfléchir à la nature phonétique possible de ces coefficients. Il fallait trouver des phonèmes capables de se vocaliser, capables d'influencer le timbre des voyelles environnantes et capables d'impacter l'articulation des consonnes. L'hébreu et l'arabe abondent en phénomènes de ce genre et on rebaptisa progressivement *laryngales* ces coefficients sonantiques. Citons ce qu'en dit Martinet (1986, p. 141) :

Les langues sémitiques illustrent bien ce type d'action de certaines consonnes sur les voyelles voisines ; comme ces consonnes y sont désignées, souvent à tort d'ailleurs, comme des « laryngales », ce terme a été vite employé pour nommer les trois différents *schwas*. Tout ce qui vient d'être exposé et ce qui va suivre forme ce qu'on appelle la « théorie des laryngales ». Pendant longtemps, cette théorie a conservé le tour largement algébrique que lui avait donné Saussure. On y opérait avec des formules comme  $*e\text{ə}_2=*â$ , sans chercher à savoir ce que pouvait être physiquement  $\text{ə}_2$  et l'on continuait à utiliser, pour noter les « laryngales », le signe vocalique  $\text{ə}$ . Puis l'habitude s'est établie d'utiliser *H* accompagné des mêmes chiffres, sans que cela entraîne nécessairement une réflexion relative à la nature des sons en question.

Oswald Szemerényi s'est intéressé à l'appropriation du concept de ces *coefficients-schwas-laryngales* au sein des comparatistes à partir de 1870<sup>18</sup>. On peut distinguer deux périodes : avant et après la découverte du hittite. Le premier protagoniste est Hermann Möller (1850-1923), un Danois, qui était un chaud partisan « de la parenté de l'indo-européen et du sémitique »<sup>19</sup>. En 1879, il appelait les laryngales « glottales, en 1880, gutturales, et on le voit continuer à hésiter entre les deux termes pendant de longues années ». En 1906, Möller écrit<sup>20</sup> : « *Es waren wahrscheinlich Gutturale von der Art der Semitischen.* » C'est en 1911 que le terme de *laryngales* fait sa première apparition : « *Die von F. de Saussure für das Vorindogermanische erschlossenen "phonèmes" entsprechen den semitischen Laryngalen.* »<sup>21</sup>

Oswald Szemerényi conclut :

Aussi pouvons-nous conclure que Saussure est bien le fondateur des vues modernes sur le vocalisme et le système des alternances apophoniques de l'indo-européen, mais n'est, au mieux, qu'un précurseur du laryngalisme ; le

---

18. Oswald Szemerényi, *La théorie des laryngales de Saussure à Kuryłowicz et à Benveniste* (BSL, n° 68, 1973). Cet article a été republié dans *Scripta Minora*.

19. D'une certaine façon, la double affirmation d'une interprétation laryngaliste et d'un apparentement au sein d'un proto-euro-hamito-sémitique a rendu la première assertion suspecte dans un premier temps. De nos jours, les indo-européanistes ont retenu la première sans la deuxième.

20. Möller (1906, Teil I (*Konsonanten*), p. VI). Nous traduisons : « Elles étaient vraisemblablement des gutturales d'un type semblable au sémitique. »

21. Szemerényi (BSL, n° 68, 1973, p. 7). Nous traduisons : « Les phonèmes postulés par F. de Saussure pour l'indo-européen correspondent aux laryngales sémitiques. »

véritable fondateur de la théorie laryngale est le savant danois Hermann Möller. (p. 11.)

C'est donc en cherchant une parenté entre l'indo-européen et le sémitique que Möller a introduit le concept de *laryngales* dans l'espace indo-européen.

Le deuxième protagoniste est Albert Cuny (1870-1947). D'abord réservé sur les thèses de Hermann Möller, il en fait une première recension sévère, puis :

Une meilleure connaissance de Möller mena Cuny à une conversion : en 1912, il publie un article de revue où non seulement il rend compte des théories de Möller (notamment de la théorie laryngale) mais va au-delà de son modèle sur plusieurs points importants<sup>22</sup>.

Néanmoins, pendant 45 ans, ces *coefficients-schwas-laryngales* restèrent une hypothèse abstraite justifiée par sa capacité à régulariser la morphologie de l'indo-européen. En 1915, le déchiffrement du hittite par Hrozny et la découverte – inattendue – qu'il s'agit d'une langue indo-européenne apportèrent une légitimation concrète à ce qui était jusque-là une audace. Bien que le hittite ne conserve pas de trace claire de toutes les laryngales postulées et que le système d'écriture soit un peu difficile<sup>23</sup>, le hittite fut interprété comme la preuve *a posteriori* que le formalisme laryngaliste était le bon.

Mais le coup le plus spectaculaire fut assurément la découverte que *h* hittite continuait la laryngale i.e.  $H_2$ , cf. *hantetsi-* premier : latin *ante, anterior*. D'autre part,  $H_1$  ne laissait pas de trace en hittite, du moins en position initiale. Cf. *estsi* « il est », de  $H_1$ es-ti<sup>24</sup>.

Avec des laryngales, le format de la morphologie est le suivant :

Données	Format laryngaliste	Format saussurien avec coefficients	Format empirique avec schwas
$\hat{e} \sim \hat{o} \sim e$	$eH_1 \sim oH_1 \sim H_1$	$eE \sim oE \sim \emptyset E$	$e\text{ə}_1 \sim o\text{ə}_1 \sim \text{ə}_1$
$\hat{a} \sim \hat{o} \sim a$	$eH_2 \sim oH_2 \sim H_2$	$eA \sim oA \sim \emptyset A$	$e\text{ə}_2 \sim o\text{ə}_2 \sim \text{ə}_2$
$\hat{o} \sim \hat{o} \sim o$	$eH_3 \sim oH_3 \sim H_3$	$eO \sim oO \sim \emptyset O$	$e\text{ə}_3 \sim o\text{ə}_3 \sim \text{ə}_3$
$st\hat{a} \sim sta$	$steH_2 \sim stH_2$	$steA \sim stA$	$ste\text{ə}_2 \sim st\text{ə}_2$
$d\hat{o} \sim do$	$deH_3 \sim dH_3$	$deO \sim dO$	$de\text{ə}_3 \sim d\text{ə}_3$

Emile Benveniste, dans *Origines de la formations des noms en indo-européen*, a assez bien résumé la situation :

La condition préalable à toute reconstruction de l'indo-européen a été fournie par la géniale découverte de F. de Saussure relative à la nature

22. *Ibidem*, p. 13. Cuny systématise les intuitions de Möller et découvre des phénomènes nouveaux, qui confortent les hypothèses initiales.

23. Voir plus loin dans l'article.

24. *Ibidem*, p. 17, à propos des travaux de Jerzy Kuryłowicz, qui est le troisième protagoniste avec Möller et Cuny.

consonantique<sup>25</sup> du phonème ə [=schwa=coefficient]. Admise et enrichie par Möller, par MM. Pedersen et Cuny, cette théorie peut aujourd'hui [1935] passer pour établie grâce à la perspicacité de M. J. Kuryłowicz, qui a su reconnaître dans le *h* hittite deux des trois variétés du ə indo-européen. (p. 148.)

La morphologie oblige donc à postuler au moins trois entités ou laryngales H, dont la nature consonantique est certaine : « *Der Primär konsonantische Charakter dieser uridg Phoneme steht ausser Frage.* »<sup>26</sup> Que nous traduisons : « Le caractère avant tout consonantique de ces phonèmes ne fait aucun doute. »

### Le système orthodoxe à trois laryngales

Si l'on prend au sérieux ces entités, il se pose la question de la nature linguistique et phonologique de ces objets que la description des langues attestées oblige à postuler dans un état de langue plus ancien. Toute la problématique du phonétisme indo-européen depuis 1870 est concentrée autour de deux questions :

- attendu que la morphologie régulière de la proto-langue amène à postuler des phonèmes \*H fantômes, quel est leur nombre ?
- une fois actée la légitimité de plusieurs phonèmes H, quels sont les meilleurs symboles pour eux dans l'API, quelle est leur nature ?

Pour démêler ces questions, plusieurs angles d'attaque sont possibles :

- la morphologie vocalique des racines, comme l'a fait Saussure,
- les alternances liées au contact consonne-laryngale, comme dans √ p\_h, « boire » > latin *bib-ere*,
- le témoignage explicite de certaines langues, anatoliennes en particulier,
- des considérations théoriques, typologiques ou phylogénétiques.

Plusieurs remarques préliminaires sont possibles. Le choix du terme *laryngale* plutôt que *coefficient* est déjà en soi un jugement de valeur qui oriente la recherche dans une direction quasi exclusive. En fait, à peu près n'importe quelle consonne peut s'amuir en allongeant la voyelle précédente. En outre, les laryngales sont définies de façon extrinsèque par leurs effets sur les voyelles et les consonnes plus que par leurs traits intrinsèques. Chaque laryngale est donc une classe de phonèmes plutôt qu'un phonème unique. C'est surtout Martinet, en bon phonologue, qui a insisté sur ce point. Et nous verrons plus loin le bien-fondé de cette remarque. La plupart des indo-européanistes fonctionnent comme si chaque H<sub>1/2/3</sub> était un phonème unique et non une classe. Et cela complique en pratique les identifications possibles. Chaque hypothèse phonétique convoque en sa faveur la partie des faits qui lui convient le mieux, en feignant d'ignorer qu'elle n'explique pas tous les faits connus et légitimes.

Il y a encore deux effets des laryngales que nous n'avons pas signalés. En sanscrit, \*/rH/ donne un /r:/ long en opposition à /r/ simple. En lituanien et en

---

25. En réalité, d'après Szemerényi, Saussure n'a jamais considéré ses propres coefficients comme des consonnes mais a toujours adhéré à une conception vocalique des coefficients.

26. Meier-Bruegger (2002, p. 106, § L314).

serbe, la séquence remontant à \*voyelle + rH-C n'est pas intonée de la même façon que la séquence remontant à \*voyelle + r-C<sup>27</sup>. De ce fait, ces langues sont tonales, ce qui est très rare en Europe.

De façon générale, la théorie actuelle la plus orthodoxe postule trois laryngales(-phonèmes). Dans son livre qui est le premier à s'intéresser de façon ciblée aux traces des laryngales en latin, Peter Schrijver indique<sup>28</sup> :

I essentially follow the views of what can nowadays be considered orthodox laryngeal theory (See e.g. Mayhofer, 1985 ; Beekes, 1988a). I shall not discuss the rich variety of alternative proposals, e.g. Adrados' palatalized and labialized laryngeals, Puhvel's nine laryngeals and Szemerényi's one laryngeal. [...] These laryngeals are written here \*h<sub>1</sub>, \*h<sub>2</sub> and \*h<sub>3</sub> (cover-symbol H). Their exact phonetic nature is unknown and is in fact irrelevant to their existence, but Indo-Europeanists agree that they were consonants.

[Nous traduisons : Je suis les idées qui peuvent maintenant être considérées comme la théorie laryngaliste orthodoxe. Je ne discuterai pas les nombreuses propositions alternatives, comme les laryngales palatalisées et labialisées de Adrados, les neuf laryngales de Puhvel ou la laryngale unique de Szemerényi. Ces laryngales sont écrites ici \*h<sub>1</sub>, \*h<sub>2</sub> et \*h<sub>3</sub> (symbole générique H). Leur nature phonétique exacte est inconnue. Cela ne met d'ailleurs pas en cause leur existence, mais les indo-européanistes s'accordent à en faire des consonnes.]

La pierre de touche de la théorie des laryngales est le témoignage du hittite et des autres langues anatoliennes : le palaïte, le louvite cunéiforme et le louvite hiéroglyphique. Depuis la découverte de Kuryłowicz, tout le monde s'accorde à voir dans le <-h(h)-> anatolien une trace directe de \*h<sub>2</sub> : « *Das Anatolische hält mit seinem h direkte Spuren von uridg \*h<sub>2</sub> fest.* »<sup>29</sup>

Il existe plusieurs exemples canoniques :

- *ha-ap-pa* « fleuve » < \*H<sub>2</sub>ep (Pokorny 51) ;
- *ha-as-te-ir-za* « étoile » < \*H<sub>2</sub>stêr (Pokorny 1027), grec *astêr* « astre » ;
- *har-ki-iš* « blanc » < \*H<sub>2</sub>erg- (Pokorny 64), grec *argês* « blanc » ;
- *pa-ah-ha-as-* « protéger, garder » < \*peH<sub>2</sub>-s (Pokorny 787), latin *pâs-tor* « pâtre, gardien ».

A contrario, \*H<sub>1</sub> n'a pas de reflet graphique dans les mots suivants :

- *ed-mi* « je mange » < \*H<sub>1</sub>ed (Pokorny 287) ;
- *eš-mi* « je suis » < \*H<sub>1</sub>es (Pokorny 340).

C'est sur cette base que la théorie laryngaliste a été définitivement validée dans ses principes. L'interprétation des graphies anatoliennes est toutefois difficile et il existe peu de certitudes encore aujourd'hui. L'opposition entre sourdes héritées et sonores n'est indiquée qu'à l'intervocalique -CC- versus -C-. Ailleurs,

27. Ces phénomènes ont été décrits dès Saussure et Cuny. Le comparatiste actuel vit sur des bases heuristiques qui ont finalement plus d'un siècle.

28. Schrijver (1991, p. 2).

29. Meier-Bruegger (2002, p. 123).

l'opposition est soit perdue diachroniquement soit non notée. Le point de vue actuel d'un spécialiste de l'anatolien est le suivant<sup>30</sup> :

In general cuneiform -vt-tv- spellings are used for inherited voiceless stops and -v-tv- for inherited voiced or voiced aspirate stops. This applies to Palaic and CLuvian, as well as Hittite. (p. 16.)

[En général, les géminées sont utilisées pour les occlusives sourdes de l'indo-européen et les simples pour les occlusives sonores et sonores aspirées. Cela vaut pour le palaite, le louvite cunéiforme et le hittite.]

The Akkadian syllabary has a series of signs for a consonant conventionally transliterated as *h*. The sound in Akkadian is apparently a voiceless velar fricative. In Hittite words *h* reflects the PIE "laryngeals"  $*h_2$  and  $*h_3$ . Orthographically, *h* patterns like the stops with contrastive *-hh-* and *-h-* between vowels, usually *-h-* in clusters (but occasionally geminate). Once again regular morphophonemic alternations such as strong stem *nâh-* versus weak *nahh-* supports the assumption that *-hh-* versus *-h-* is contrastive. [...] Historically, geminate *-hh-* is the regular reflex of  $*h_2$ , while all clear cases of medial *-h-* continue "lenited"  $*h_2$ . (p. 21.)

[Le syllabaire akkadien a une série de signes pour une consonne conventionnellement transcrite <h>. Le son akkadien est apparemment une fricative vélaire sourde. Dans les mots hittites *h* reflète les « laryngales »  $*h_2$  and  $*h_3$ . Orthographiquement, *h* se comporte comme les occlusives : *-hh-* et *-h-* s'opposent entre voyelles, *-h-* est simple dans les groupes de consonnes (mais parfois géminée). Répétons que les alternances morphologiques telle que la racine forte *nâh-* versus faible *nahh-* conforte l'hypothèse que *-hh-* s'oppose à *-h-*[...] La géminée *-hh-* est le reflet régulier de  $*h_2$  alors que tous les exemples clairs de *-h-* intervocalique continue  $*h_2$  « lénifiée ».]

My assumption to pharyngeal articulation for Proto-Anatolian and the cuneiform languages is not crucial, and velar fricatives instead are quite possible. (p. 22.)

[Mon hypothèse d'une articulation pharyngale pour le proto-anatolien et les langues cunéiformes n'est pas cruciale, et des fricatives vélares sont tout à fait possibles.]

$/*h_1/$  I know of no compelling evidence for the preservation of PIE  $/*h_1/$  in Proto-Anatolian in any position. [...] In most positions it is lost without a trace. (p. 65.)

[Je ne connais pas d'exemple définitif de la conservation de  $/*h_1/$  en proto-anatolien en aucune position. [...] Elle disparaît sans une trace dans la plupart des positions.]

$/*h_2/$  PIE  $/*h_2/$  is generally preserved in Proto-Anatolian as a fortis, voiceless fricative which I symbolize as H. (p. 68.)

[ $/*h_2/$  est généralement conservée en proto-anatolien comme une fricative forte sourde, que je symbolise <H>.]

---

30. Melchert (1994). Voir également Kimball (1999). Leurs points de vue sont proches, ce que Melchert reconnaît lui-même dans la préface de son livre. Entre crochets, notre traduction.

*/\*h<sub>3</sub>/* The only major controversy regarding laryngeals in Anatolian concerns the fate of initial */\*h<sub>3</sub>/*. (p. 71.) [...] */\*h<sub>3</sub>/* is preserved initially as *h-* in Hittite, Palaic and Cuneiform Luvian. I assume that initial */\*h<sub>3</sub>/* was a lenis voiced fricative */\*h/* in Proto-Anatolian, distinct from the fortis, voiceless fricative */\*H/* which is the regular reflex of */\*h<sub>2</sub>/*. (p. 72.)

[La seule controverse majeure concernant les laryngales en anatolien concerne le sort de */\*h<sub>3</sub>/*. [...] */\*h<sub>3</sub>/* est conservée à l'initiale comme #h- en hittite, palaite et louvite cunéiforme. Je suppose que */\*h<sub>3</sub>/* à l'initiale était une fricative voisée douce */\*h/* en proto-anatolien, distincte de la fricative sourde forte */H/* qui est le reflet régulier de */\*h<sub>2</sub>/*.]

Melchert considère que *\*H<sub>1</sub>* ne laisse aucune trace et que *\*H<sub>2</sub>* est sans doute une fricative sourde vélaire */x/* ou pharyngale */H/*, proche du <h> de l'akkadien.

Le tableau suivant résume différentes hypothèses de système à trois laryngales(-phonèmes) :

	H <sub>1</sub>	H <sub>2</sub>	H <sub>3</sub>
Melchert <sup>31</sup>	(∅)	/x/ ou /H/	/h/
Meier-Brügger <sup>32</sup>	/h/	/x/	/ɣ <sup>(w)</sup> /
Beekes, Schrijver <sup>33</sup>	/?/	/ɣ/	/ɣ <sup>w</sup> /

Le constat émergeant de cette comparaison est un désaccord apparent sur le lieu – glottal, pharyngal, vélaire – et le mode – sourd, sonore, glottalisé, labialisé. La reconstruction interne des laryngales dans un cadre strictement indo-européen se heurte à des difficultés sérieuses. Dans les faits, il existe plusieurs systèmes orthodoxes à trois laryngales (-phonèmes).

Néanmoins, il faut noter trois points de convergence générale :

1) H<sub>3</sub> est unanimement tenue pour sonore.

En effet la racine *\*p-H<sub>3</sub>* « boire » fournit le corpus : grec *pô-tis* < *\*p-H<sub>3</sub>* « boisson », sanscrit *pibati* « il boit » < *\*pipH<sub>3</sub>-ati* dans lequel *\*p* est sonorisé par *\*H<sub>3</sub>*, latin *bib-ere* « boire » < *\*pib-* < *\*pipH<sub>3</sub>-*.

Dans le tableau, H<sub>3</sub> est toujours sonore mais de lieu variable : glottal, pharyngal, vélaire.

2) H<sub>3</sub> est considérée de même lieu articulaire que H<sub>2</sub> avec un effet labialisant en supplément.

3) H<sub>1</sub> est censée avoir un autre lieu articulaire que le couple H<sub>2</sub>/H<sub>3</sub>.

Le système proposé par Melchert est un peu différent car il concerne un état de langue déjà évolué par rapport à l'indo-européen et valable pour le proto-anatolien.

31. Les valeurs proposées par Melchert concernent le proto-anatolien. Conforme à sa ligne prudente, Melchert ne propose pas de valeurs pour l'indo-européen. Il met même des guillemets au mot « laryngeal ».

32. Meier-Bruegger (2002, p. 106).

33. Schrijver (1991, p. 2).



### Les données hétérodoxes

L'orthodoxie à trois laryngales(-phonèmes) est motivée en premier lieu par la volonté, *sans aucun doute légitime en l'absence de certitudes*, de limiter le nombre de phonèmes disparus. Néanmoins, cela aboutit en pratique à mutiler les données de façon incontestable.

Les spécialistes de l'albanais postulent une quatrième laryngale, dont il est rarement fait état. Par exemple, comparé au grec *orkhis* « testicule », l'albanais présente *herðe*<sup>34</sup>. Cette langue a donc des traces explicites, encore aujourd'hui, de laryngales. *A contrario*, les aspirées de l'arménien peuvent s'expliquer intégralement comme un développement interne tardif, ce qui ruine toute possibilité de s'appuyer sur cette langue.

D'autre part, les données anatoliennes ne sont pas aussi orthodoxes qu'on veut bien le dire :

- *ha-ap-pa* « fleuve » < \*H<sub>2</sub>ep (Pokorny 51 \*ap) ;
- *ap-pa* « derrière » < \*H<sub>2</sub>ep (Pokorny 53 \*apo)<sup>35</sup> ;
- *me-hur* « temps » < \*meH<sub>1</sub> (Pokorny 703 \*mê) ;
- *me-e-hu-e-ni* « temps » < \*meH<sub>1</sub> (Pokorny 703 \*mê)<sup>36</sup>.

Il existe donc des exemples clairs où H<sub>2</sub> n'a aucun reflet et, *a contrario*, où H<sub>1</sub> est reflété.

Jaan Puhvel, un autre spécialiste du hittite, postule une double valeur phonétique pour chaque laryngale, suivant qu'elle laisse ou non une trace <h> en hittite<sup>37</sup> :

- E<sub>1</sub> = voiceless e-coloring laryngeal, lost in Hittite, intervocalically lengthens preceding vowel and yields glide -y- ; E<sub>1</sub> > a. [Équivaut à H<sub>1</sub> orthodoxe.]
- E<sub>2</sub> = voiced e-coloring laryngeal, Hittite *h-*, *-h-*.
- A<sub>1</sub> = voiceless a-coloring laryngeal, Hittite *h-*, *-h(h)-*. [Équivaut à H<sub>2</sub> orthodoxe.]
- A<sub>2</sub> = voiced a-coloring laryngeal, lost in Hittite, A<sub>2</sub> > a.
- A<sub>1</sub><sup>w</sup> = voiceless o-coloring laryngeal, lost in Hittite, A<sub>1</sub><sup>w</sup> > u.
- A<sub>2</sub><sup>w</sup> = voiced a-coloring laryngeal, lost in Hittite, Hittite *h-*, *-h-*. [Équivaut à H<sub>3</sub> orthodoxe.]

Cette analyse plus complexe que celle de Melchert ou Kimball est donc induite par les données, qui ne se laissent pas réduire à une orthodoxie à trois laryngales(-phonèmes) seulement.

Un autre exemple de Puhvel qui n'est pas dans Pokorny est le suivant : hittite *ay-*, *e-* « être chaud » < \*i.e. H<sub>2</sub>ei-dh « brûler, être en flammes, être chaud » (Pokorny 11 \*ai-dh). Puhvel rapproche le mot hittite de l'albanais *hî* « cendre ».

---

34. Mallory (1997b, p. 10). Indiqué également dans Pokorny (p. 782) sans mention particulière. Voir hittite *ark* « pénis », sans #h-.

35. Cité par Jerzy Kuryłowicz, le découvreur des laryngales, dans Kuryłowicz (1956, p. 225).

36. Lehmann (1955, p. 26).

37. Puhvel (1984, vol. 1, p. X).

Le *Hethitisches Elementarbuch* de Johannes Friedrich a été traduit en anglais, remanié et mis sur Internet par Olivier Lauffenburger. Dans la traduction anglaise *Hittite Grammar* (p. 16), le traducteur écrit :

In the most common theory, P.I.E. had three laryngeals, noted  $H_1$ ,  $H_2$  and  $H_3$  that could “color” a neighboring vowel ‘e’. The laryngeal  $H_1$  had no coloration effect, the laryngeal  $H_2$  colored in ‘a’ and the laryngeal  $H_3$  colored in ‘o’. In Hittite, the laryngeal  $H_1$  vanished and the laryngeal  $H_3$  was retained only in initial position. In median position, the fricative resulting from a laryngeal can be lenis (written between two vowels by ‘h’) or fortis (written between two vowels by ‘hh’). [...] It should be noted that the theory described here is incomplete : it does not explain cases where Hittite displays a ‘h’ where there is no laryngeal, and conversely cases where Hittite does not display a ‘h’ where a laryngeal occurred.

[Dans la théorie la plus répandue, l’indo-européen avait trois laryngales, notées  $H_1$ ,  $H_2$  et  $H_3$ , qui pouvaient « colorer » une voyelle « e » voisine. La laryngale  $H_1$  n’avait pas d’effet colorant, la laryngale  $H_2$  colorait en « a » et la laryngale  $H_3$  colorait en « o ». En hittite, la laryngale  $H_1$  a disparu et la laryngale  $H_3$  est conservée seulement en position initiale. En position médiane, la fricative résultant d’une laryngale peut être douce (*lenis*) (écrite entre voyelles « h ») ou forte (*fortis*) (écrite entre voyelles « hh »). [...] Il faut noter que la théorie décrite ici est incomplète [*sic*] : elle n’explique pas les cas où le hittite présente un « h » alors qu’il n’y a pas de laryngale<sup>38</sup> et à l’inverse les cas où le hittite ne présente pas de « h » alors qu’une laryngale existait.]

Si l’on accorde crédit aux données hétérodoxes, l’opposition entre  $H_1$  et  $H_2/H_3$  n’est pas une opposition de *lieu* mais une opposition de *mode*. Il existe au minimum deux couples de fricatives  $H_2/H_3$  *sonores* et il en résulte deux fricatives  $H_1$  *sourdes*, en opposition phonologique avec  $H_2/H_3$ , à savoir partageant le même lieu articulatoire et s’opposant par le seul trait : sonore ~ sourde. D’autre part, si l’on accorde crédit au système glottalique, /ʔ/ existe aussi, en tant que primitive phonologique.

Il existe une hypothèse remontant simultanément à Edward Sapir et à Antoine Meillet que les laryngales  $H_1$  et  $H_2$  puissent se durcir pour donner -k-. Un bon exemple est le féminin régulier latin *imperâtor*, formé avec le suffixe régulier -i $H_1$ -, qui se dit *imperâtri-k-s* au nominatif singulier (-s#). La racine \*gwi $H_3$  *vivant* donne *quick* en anglais, dans lequel la finale -ck n’est sans doute pas un suffixe mais le durcissement de  $H_3$ . Il a donc de fortes raisons de penser qu’à côté des glottales et des pharyngales, il a existé une série vélaire/uvulaire /x/ /ɣ/ /ɣ<sup>w</sup>/. Le hittite fournit des indications internes dans ce sens. A côté des formes *iskisa* « dos », *tetkissar* « tempête » et *hameskanza* « printemps », il existe des variantes *ishisa*, *tethessar* et *hameshanza*. En outre, le nom de personne *Giluhepa*, d’origine hourrite, est écrit <Krgp> en égyptien hiéroglyphique. La vélarité de certaines réalisations « laryngales » est donc indéniable d’après un faisceau de données internes et externes.

---

38. Nous ne connaissons pas de cas de ce type et l’auteur n’en donne pas.

Au final, il existe un ensemble convergent de constats :

- des alternances internes au hittite,
- des transcriptions en langue étrangère,
- des correspondances duales à l'intérieur du corpus indo-européen.

La théorie orthodoxe limitant à trois valeurs phonétiques les *laryngales* nous paraît intenable. Les données amènent à postuler des systèmes plus riches. Le proto-système, que nous proposons pour l'indo-européen, présente au minimum l'inventaire phonologique suivant :

Mode/Lieu	Glottale	Pharyngale	Vélaire/Uvulaire
Sourd		H <sub>1,b</sub> */H/	H <sub>1,c</sub> */x/
Sonore		H <sub>2,b</sub> */ʕ/	H <sub>2,c</sub> */ɣ/
Glottalisé	H <sub>0,a</sub> */ʔ/		
Labialisé		H <sub>3,b</sub> */ʕʷ/	H <sub>3,c</sub> */ɣʷ/

Dans l'état actuel du dossier indo-européen, il nous semble qu'on ne peut plus progresser sur des bases internes. Nous allons maintenant examiner les données extérieures, macro-comparatives, en particulier les données sémitiques.

### Les données macro-comparatives

Commençons par H<sub>1</sub> qui a deux valeurs possibles : /ħ/ ח et /h/ ח.

Racine i.e. Pokorny 703 \*mê √ m\_H<sub>1</sub> *mesurer (entre autres le temps)* (moment, mois, année).

Racine i.e. Pokorny 731 \*mên-ôt √ m\_H<sub>1</sub>-n *mois, lunaison, lune.*

- Gotique *mel* [me :l] « temps »
- Lituanien *métas* « temps, année »
- Hittite *me-hur* « temps »
- Hittite *me-e-hu-e-ni* [meħweni] « temps »
- Latin *mensis*, Grec *mênê* « mois »
- Anglais *moon, month* « lune, mois »
- Arabe *maħwaʔ* « heure, moment »
- Hébreu *maħzor* « cycle, révolution (des astres) »
- Hébreu *maħzor hayareaħ* « lunaison (cycle de la lune) »
- Cette racine serait √ m\_h<sub>1,b</sub> [m\_h] avec une pharyngale.

Racine i.e. Pokorny 666 \*lê(i) √ l\_H<sub>1</sub> *faible, mou, lent, en retard*

- Latin *lênis* « doux ; calme, lent »
- Lituanien *lenas* « calme, lent »
- Lette *léns* « paresseux, doux »
- Vieux slave *lenu* « indolent »
- Latin *lentus* « indolent ; lent, long »
- Grec *elînuô* « se reposer »
- Arabe *lahħ* « lent, paresseux, rétif (un animal) »
- Arabe *lahlah* « rester à sa place et ne pas bouger »
- Arabe *laham* « ne plus vouloir avancer »
- Arabe *lahamaʔ* « lent et paresseux » (variante avec x au lieu de h)

Latin *lentus* « mou, souple »  
 Arabe *lahīʿ* « être lâche et flasque (chairs) »  
 Arabe *lahmaʿ* « langueur et flacidité du corps »  
 Latin *lassus* « fatigué »  
 Arabe *laqab* « être très fatigué »  
 Il y aurait trois racines différentes : \*√ l\_H<sub>1,b</sub> \*√ l\_H<sub>1,c</sub> \*√ l\_H<sub>2,c</sub>

Racine i.e. Pokorny 661 \*lei-(bh) *faible, affamé*

Grec *lîmos* « faim » < \*liH-mos<sup>39</sup>  
 Arabe *latah* « avoir faim »  
 Arabe *lathân, lathân* « affamé »  
 Arabe *mulhûb* « décharné »  
 Arabe *malhûs* « décharné »  
 Arabe *lahab* « être très maigre (de vieillesse) »  
 Arabe *lahiq* « être mince »  
 Il y a une seule racine : \*√ l\_H<sub>1,b</sub>.

Racine i.e. Pokorny 662 \*s-lei- *mucus, boue*

Latin *lentus* « visqueux, tenace »  
 Anglais *slime* « mucus » < \*s-lî-m  
 Allemand *Schlei* « tanche » (cf. balto-slave)  
 Arabe *lahh* « avoir les paupières collées »  
 Arabe *lahij* « être couvert de saletés visqueuses (paupières, yeux) » (Variante avec *h*)  
 Arabe *multahimaʿ* « conjonctivite »  
 Hébreu *lahmît* « conjonctivite »  
 Il y a une seule racine : \*√ l\_H<sub>1,b</sub>.

Racine i.e. Pokorny 683 \*lêu- *chant*

Latin *laud* « éloge »  
 Allemand *Lied* « chant » < \*leH-ut  
 Arabe *lahn* « son, mélodie, chant »  
 Hébreu *lahan* « air, mélodie »  
 Il y a une seule racine : \*√ l\_H<sub>1,b</sub>.  
 Voir plus loin le nom de l'alouette.

Racine i.e. Pokorny 662 \*lei- /\*leu- *enduire, oindre ; souiller, tacher*

Latin *lînô* « oindre »  
 Irlandais *as-len-aim* « je souille »  
 Grec *lûma* « souillure, ordure »  
 Arabe *lahh* « oindre, imprégner »  
 Hébreu *lihîlûh* « saleté, crasse »  
 Hébreu *lihleh* « salir, souiller »  
 Il y aurait 1 seule racine : \*√ l\_H<sub>1,c</sub>.

39. Ce schème apophonique i\_o est inhabituel en indo-européen.

Racine  $\sqrt{H_{1,b}_d}$  un

Slave \**od* \**ed* « un »

Arabe *'ahad* / *waḥid* « un »

Les exemples avec  $H_2$  qui a trois valeurs possibles : /' /  $\acute{a}$  , /' /  $\varepsilon$  et / $\acute{g}$  /  $\acute{e}$ .

Racine i.e. Pokorny 4 ag  $\sqrt{H_{0,a}_g}$  mener, guider

Latin *agō* « pousser devant soi, faire avancer »

Latin *agmen* « troupe d'hommes »

Arabe *'ajā'* « forcer, contraindre à quelque chose »

Arabe *'ajl* « troupeau ; troupe d'hommes »

Racine i.e. Pokorny 38 anH  $\sqrt{H_{0,a}_n H_{1,b}}$  souffler, vent, respirer

Latin *anima* « souffle ; âme »

Grec *anemos* « vent »

Arabe *'anaḥ* « haleter, respirer avec effort ; pousser un soupir »

Hébreu *'anaḥah* « soupir »

Racine i.e. Pokorny aner  $\sqrt{H_{0,a}_n r}$  homme

Grec *anēr* « homme »

Arabe *'ins* « homme, être humain »

Hébreu *'anš* « homme »

Racine i.e. Pokorny 1010 s-tâ  $\sqrt{t_{H_{2,b}}}$  couler, s'écouler

Grec *sta-zō*, *sta-lazzō* « s'écouler »<sup>40</sup>

Arabe *tā'-b* « faire écouler, verser (larmes, sang) »

Arabe *tā'-jar* « verser, répandre »

Racine s-tâ  $\sqrt{t_{H_{2,c}}}$  mare, étang

Latin *stāgnum* « étang »

Arabe *tāg-b* « mare d'eau stagnante »

Valeurs supplémentaires où  $H_{2,a}$  est /h/ ◊

Racine aw  $\sqrt{H_{2,a}_w}$  air, souffle

i.e. aw  $\sqrt{H_{2,a}_w}$  « air »

Arabe *hawa'* « air, souffle »

Racine ay  $\sqrt{H_{2,a}_y}$  vivant

i.e. ay  $\sqrt{H_{2,a}_y}$  « vivant »

i.e. ayənk-  $\sqrt{H_{2,a}_y}$  « jeune, plein de vie »

Arabe *haya'* « vivre »

Valeurs supplémentaires où  $H_{2,a}$  /h/ ◊ alterne avec  $H_{2,c}$  / $\acute{g}$ /

---

40. Dans plusieurs langues indo-européennes (celte, balto-slave, germanique), cette racine donne le nom de l'urine. Cette racine semble être associée aux différents fluides corporels (larmes, sang, etc.). Peut-être est-ce un mot très ancien du domaine « médical ».

Racine i.e. Pokorny 1053 tâ √ t<sub>H<sub>2,a</sub></sub> *fondre, se liquéfier*

Latin *tâ-bêo* « fondre (neige, cire) »

Anglais *to thaw* « fondre (neige, glace) »

Arabe *taḥṭah* « fondre (se dit de la neige) »

Arabe *taḡab* « fonte des glaces »

Si l'on accepte ces rapprochements, l'enveloppe des réalisations phonétiques, que nous proposons pour l'indo-européen, présente donc au minimum l'inventaire suivant :

Mode/Lieu	Glottale	Pharyngale	Vélaire/Uvulaire
Sourd		H <sub>1,b</sub> *[H]	H <sub>1,c</sub> *[x]
Sonore	H <sub>2,a</sub> *[h]	H <sub>2,b</sub> */'/	H <sub>2,c</sub> *[ɣ]
Glottalisé	H <sub>0,a</sub> */?/		
Labialisé		H <sub>3,b</sub> */ <sup>w</sup> /	H <sub>3,c</sub> */ɣ <sup>w</sup> /

Certains phonèmes ne sont pas complètement indépendants :

– H<sub>2,a</sub> \*[h] et H<sub>2,c</sub> \*[ɣ] alternent,

– H<sub>1,b</sub> \*[H] et H<sub>1,c</sub> \*[x] alternent.

En sémitique, et en particulier en arabe, il s'agit de phonèmes différents. La question reste ouverte de savoir si ces paires de phonèmes alternants peuvent être ramenés à l'unité. A un stade linguistique plus ancien que l'indo-européen ou le proto-sémitique, il est possible que chaque paire de phonèmes alternants corresponde à un proto-phonème unique, mais les clés de distribution allophonique qui pourraient valider cette hypothèse de réduction à l'unité nous échappent pour le moment.

### La structure de la racine

L'influence du sémitique sur la reconstruction de l'indo-européen ne s'arrête pas à la phonologie. Rappelons en préalable qu'Emile Benveniste était né en 1902 à Alep, en Syrie. A nos yeux, cette origine proche-orientale est la raison directe de son intérêt pour la forme des racines indo-européennes et elle explique aussi le type de format proposé. Il ne fait quasiment jamais allusion au sujet – plus ou moins tabou – du sémitique mais dans les faits, ce qu'il propose aligne l'indo-européen sur l'écriture consonantique de l'arabe et des langues sémitiques.

Ses objectifs sont présentés dans la préface de son livre *Origines de la formation des noms en indo-européen* :

L'objet essentiel de la grammaire comparée, depuis une soixantaine d'années [de 1870 à 1935], a été de poser des correspondances entre les langues indo-européennes et d'expliquer, en partant de l'état que définissent ces correspondances, le développement des dialectes attestés [= les langues indo-européennes elles-mêmes]. [...] Depuis le *Mémoire* de F. de Saussure, le problème de la structure des formes indo-européennes elles-mêmes a été

presque complètement négligé. Il paraît communément reçu qu'on peut analyser l'évolution de l'indo-européen sans se soucier de ses origines, qu'on peut comprendre des résultats sans pousser jusqu'aux principes. De fait, on ne va guère au delà de la constatation. [...] Nous avons visé avant tout à définir des structures, des alternances, l'appareil formel.

Cet état des lieux rédigé en 1935 sur l'empirisme du domaine indo-européen au ras des données lexicales reste largement vrai en 2007. Plus loin, il précise sa pensée :

Ce qu'on a enseigné jusqu'ici de la nature et des modalités de la racine est, au vrai, un assemblage hétéroclite de notions empiriques, de recettes provisoires, de formes archaïques et récentes, le tout d'une irrégularité et d'une complication qui défient l'ordonnance. On enregistre des racines monosyllabiques (\*bher-) ou dissyllabiques (\*g<sup>w</sup>eyə-); des racines bilitères (\*dô-), trilitères (\*per-), quadrilitères (\*leuk-), quinquilitères (\*sneig<sup>w</sup>h-); des racines à voyelle intérieure (\*men-) ou à diphtongue (\*peik-); à voyelle initiale (\*ar-) ou à voyelle finale (\*pô-); à degré long (\*sêd-) ou à degré zéro (\*dhək-); à diphtongue longue (\*srêig-) ou brève (\*bheudh-). (p. 147.)

Dans son livre, Benveniste ne se comporte pas en comparatiste ou en reconstruc-teur. Il parle en grammairien normatif outré par le désordre. La modélisation des racines qu'il propose est la suivante :

- thème I	C <sub>1</sub> vC <sub>2</sub> (-C <sub>3</sub> -)	[fa'(-l-)]
- thème II	C <sub>1</sub> C <sub>2</sub> vC <sub>3</sub> -	[f'al-]

Les racines du type \*ar sont reformatées \*H<sub>2</sub>er-, ce qui explique que \*a ne puisse jamais alterner avec \*e à l'initiale, puisqu'il s'agit de deux allophones du même phonème, et les racines du type \*dô- sont reformatées \*deH<sub>3</sub>- : pour la même raison d'allophonie, \*ê et \*ô ne peuvent pas alterner.

Bien qu'il s'en défende<sup>41</sup>, sa pensée suit un chemin qui est strictement l'inverse de celui qui mène de l'alphabet phénicien à l'alphabet grec. En pratique, Benveniste retrace la Méditerranée dans l'autre sens sans le dire.

Cette définition [de la racine] doit être entendue littéralement et phonétiquement, et non pas au sens où les sémitistes l'emploient pour caractériser seulement le schème consonantique de la racine. (p. 171.)

Le thème I de la racine benvenistienne est en termes sémitiques la racine consonantique dotée du schème vocalique -v(-Ø-) [fa'l], et le thème II a le schème -Ø-v- [f'al]. Curieusement, Benveniste nie une filiation conceptuelle, qui est pourtant évidente.

On peut y voir une précaution pour ne pas froisser les comparatistes. Mais il y a aussi une raison technique, interne. Benveniste ne parle pas de prosodie, mais ce qui distingue le thème I [fa'l] du thème II [f'al], c'est la place de l'accent à très haute époque en indo-européen ancien. Le squelette consonantique de la racine indo-européenne n'admet qu'une seule voyelle obligatoirement accentuée, soit dans la première, soit dans la deuxième syllabe.

41. Il appartient à l'histoire des sciences de comprendre pourquoi.

Benveniste veut rendre compte en même temps de deux phénomènes différents : la structure des racines, qui sont constituées de consonnes uniquement, et la place de l'accent, qui conditionne l'insertion de la seule voyelle admissible dans la racine en indo-européen ancien. Notre avis est que les schèmes pléovocaliques du type fa'al, fa'il ou fa'ul sont impossibles dans le stade le plus ancien de l'indo-européen. Seuls sont possibles fa'l, f'al ou f'la. Cette particularité identifie les formes réellement anciennes et les schèmes pléovocaliques sont des innovations réalisées par les langues filles<sup>42</sup>.

A ce titre, on peut noter que certaines langues indo-européennes, comme le groupe germanique, et en particulier l'anglais, résistent à la pléovocalie et ont même tendance, dans les emprunts, à expurger le maximum de voyelles pleines. L'anglais perpétue un commandement phonologique hérité de l'indo-européen ancien : le schème vocalique de la racine aura une voyelle pleine et une seule, le reste sera du lubrifiant phonétique.

La théorie de Benveniste est la fusion de la réalité sémitique, consonantique, de la racine avec la prosodie, monovocalique, spécifique de l'indo-européen ancien. Elle propose d'intégrer toutes les racines dans un format unique : un squelette  $C_1-C_2-(C_3)$  + une seule voyelle.

Au terme de cette première partie, il apparaît nettement que tant la phonologie que les racines de l'indo-européen ont été modelées d'après l'exemple implicite ou explicite des langues sémitiques, et cela serait encore plus flagrant si l'indo-européen était écrit avec l'alphabet arabe plutôt que latin.

### **L'incrémentation en indo-européen**

Nous allons maintenant aborder nos propres travaux sur l'indo-européen. Par des chemins différents, nous avons été amené à une évolution de la théorie de la racine indo-européenne qui nous rapproche des principes *Racines et incréments*, tels que les propose Georges Bohas pour l'arabe. Bien que la théorie proposée par Benveniste soit un progrès évident, la fréquentation des études indo-européennes nous a convaincu depuis de nombreuses années de la nécessité d'une réforme. Mais nous avons mis très longtemps à trouver une solution acceptable. Etant donné que nous sommes à la fois (macro-)comparatiste et reconstituteur, à nos yeux, ce type de théorie permet de mettre en évidence de la *morphologie fossile*. C'est le concept par lequel nous sommes arrivé à une théorie du type *Racines et incréments*.

Dans un autre domaine, la doxa officielle de l'arabe prône une théorie de la racine à trois consonnes. En pratique, cette doxa empêche de percevoir la structure réelle du lexique, qui se trouve atomisé en une myriade de racines de même sens mais sans lien supposé. Il nous semble que le point de départ de la refondation proposée par Georges Bohas pour l'arabe est en premier lieu un refus de l'atomisation lexicale, un refus de la déstructuration du lexique et la volonté de faire émerger l'innervation qui traverse et organise le corpus lexical de l'arabe.

---

42. Précisons, pour éviter toute ambiguïté, que ce diagnostic est personnel.



Notre démarche est similaire mais sur un corpus différent : celui des langues indo-européennes. Nous partageons ce refus de l'atomisation et de la déstructuration du lexique.

Le point de départ de notre questionnement, depuis une vingtaine d'années, est un triple constat et une conviction intime :

- constat que le dictionnaire étymologique de l'indo-européen de Pokorny offre de multiples entrées de forme proche et de sens proche, qui seraient donc à considérer comme autant et autant de « racines » synonymes, de sens souvent assez vague ;
- constat que souvent, ces « racines » n'ont pas de classe syntaxique clairement explicitée ;
- constat que ce dictionnaire recourt à un concept d'« élargissement » des « racines », qui donne l'impression que n'importe quel phonème de la langue peut être un élargissement et que ces « élargissements » ne sont motivés par aucun trait de sens, ni par aucune considération de classe syntaxique ;
- conviction intime que si l'indo-européen doit être considéré comme une vraie langue, c'est-à-dire comme un état synchronique possible, et non pas comme un mécano lexico-graphique commode mais fictif, alors il est impossible que les locuteurs aient eu à leur disposition tant et tant de racines synonymes de sens vague, élargissables *ad libitum* par n'importe quel phonème sans aucune considération sémantique ni syntaxique. Il est bien connu que les langues se distinguent autant par ce qu'elles permettent que par ce qu'elles ne permettent pas. Ici, tout paraît possible.

Un exemple de mécano lexicographique est \*(s)-ker « couper », dans Pokorny (1959, p. 938-947). L'entrée elle-même donne le ton : (s)ker-, (s)kerə-, (s)kre:-. Trois formes « de base » sont annoncées avant élargissement ou suffixation, avec ou sans #s-. Indépendamment de la cohérence sémantique du corpus réuni sous cette entrée, qui fait aussi problème, l'inventaire – difficile à réaliser – est (au moins) le suivant :

- (s)-ker-, (s)-kr-ek, (s)-ker-s,
- (s)-kr-eH<sub>1</sub>, (s)-kr-eH<sub>1</sub>-t, (s)-kr-e-n-H<sub>1</sub>-d, (s)-kr-eH<sub>1</sub>-bh, (s)-kr-eH<sub>1</sub>-m,
- (s)-kr-eH<sub>2</sub>, (s)-kr-eH<sub>2</sub>-k,
- (s)-krew, (s)-krew-d, (s)-keru-p, (s)-krew-p,
- (s)-ker-d, (s)-ker-t,
- (s)-ker-b, (s)-ker-bh, (s)-kr-eb, (s)-kr-ebh, (s)-ker-p, (s)-kr-ep,
- (s)-ker-i-H<sub>1</sub>, (s)-kr-iH<sub>1</sub>, (s)-kr-eH<sub>1</sub>-y-, (s)-kr-iH<sub>1</sub>-p, (s)-ker-iH<sub>1</sub>-bh,
- (s)-kr-ey-d, (s)-kr-ey-t.

Dans ce kaléidoscope morphophonologique, on ne sait plus si on a affaire à des voyelles, des consonnes, des suffixes, des élargissements, etc. Nous sommes convaincu que jamais aucune proto-langue ne peut avoir fonctionné de cette façon. Tout cela se situe largement au-delà des limites de la linguistique.

Il nous a toujours semblé nécessaire de repenser la théorie de la racine, afin que l'indo-européen puisse *devenir* une vraie langue. Mais le chemin est long entre les premiers doutes, l'intention et la solution.

Pour des raisons quasiment idéologiques, les comparatistes du XIX<sup>e</sup> siècle ont exclu que l'indo-européen aurait pu posséder des infixes ou des préfixes. Leurs préjugés quant à une prétendue évolution hiérarchisée des langues ne l'admettaient pas. L'indo-européen se devait d'être flexionnel et suffixal, car c'était à leurs yeux le fonctionnement le plus noble et donc le plus adapté à la proto-langue indo-européenne, supposée être « supérieure ». C'est Schleicher qui est à l'origine de cette typologie hiérarchisée des langues. Les continuateurs, quelle que soit la pureté de leurs intentions et de leurs travaux, n'ont pas remis en cause ce genre de prémisses, qui ont pourtant orienté l'indo-européen dans la mauvaise direction. Avec le temps, les aspects scandaleux de la doxa indo-européenne (*aryenne...*) ont été gommés, mais les conséquences indirectes restent présentes dans les travaux contemporains.

Le seul préfixe toléré est le #s- mobile initial, dont la présence est trop évidente pour ne pas être vue, et le seul infixe toléré est -n- à valeur de présent, qui est vivant en latin. On lui doit « rompre » vis-à-vis de « rupture » (avec et sans nasale). Dans la suite, nous allons voir que tout porte à croire que l'indo-européen a possédé un système d'affixes, comprenant préfixes, infixes et suffixes. Dans un souci de convergence terminologique avec la *TME*, nous parlons désormais d'*incréments*.

Au début du XX<sup>e</sup> siècle, Hermann Möller avait lui aussi conclu à la nécessité d'une théorie des affixes. Il avait recouru au terme de (pré-)formatif.

Même dans le cadre standard, restrictif avec seulement #n- et -n-, certains mots restent isolés dans le comparatisme classique. Ainsi les différents noms du *scarabée* en grec :

Grec *kârabos* « scarabée »

Grec *kerambuks* « scarabée » (infixe -n-)

Grec *skarabeis* « scarabée (préfixe #s-) »<sup>43</sup>

Aussi étonnant que cela puisse paraître, spécialement pour quelqu'un qui s'intéresse de près à la *TME* en arabe, ces mots ne sont pas censés avoir de rapport entre eux. Il nous semble pourtant que ce rapport est évident, même si les schèmes vocaliques internes sont anormaux, vu de l'indo-européen orthodoxe. Comment ne pas voir que ces mots partagent la même racine  $\sqrt{k\_r\_b}$  ?

Notre conviction, depuis longtemps et aujourd'hui plus que jamais, est qu'on peut appliquer au corpus indo-européen la même démarche que celle qui est engagée pour le lexique de la langue arabe : à savoir déterminer des racines et des *incréments*. Dès lors qu'on fait sauter le verrou intellectuel interdisant de chercher des *incréments*, la théorie de la racine indo-européenne devient lumi-

---

43. Ces mots sont souvent tenus pour non indo-européens, mais le simple fait qu'il existe une forme avec #s- préfixé suffit à en faire une racine indo-européenne.

neusement simple et quantité de lexèmes qui ne sont pas rattachés à une racine trouvent naturellement leur place.

### Les incréments -r- et -l- en indo-européen

Outre le #s- initial et l'infixe -n-, deux incréments fréquents de l'indo-européen sont -r- et -l- qui sont toujours infixés. Les exemples sont innombrables et touchent tous les champs sémantiques. Il faut noter que contrairement à l'arabe qui tend à incrémenter -r- et -l- en toute position, l'indo-européen est restrictif et n'incrémente qu'en position infixale. En outre, il est fréquent que -r- ait une nuance péjorative.

Un premier exemple est l'anglais *to speak* « parler » vis-à-vis de l'allemand *sprechen*. L'anglais est la seule langue germanique à n'avoir aucune trace de -r- dans cette racine. Ce point a déjà été observé et discuté :

On est porté à admettre la chute<sup>44</sup> indo-européenne de r, sous certaines conditions, dans le groupe initial consonne + r, par exemple, latin *fungor* : *fruor fruges*, gotique *brukjan* « avoir besoin, se servir de » ; sanscrit *bhanakti* « briser », arménien *bekanim* « je brise » ; latin *frango*, gotique *brika* « je brise » ; grec (w)agnumi : (w)râgnumi « je brise » : néerlandais *wrak* « débris » ; grec *poti* : *proti* « à côté de » ; anglo-saxon *specan* « parler » ; vieux-haut-allemand *sprehhan*. Sur cette question, voir Brugmann, Grdr<sup>45</sup> I<sup>2</sup> 426 et ses références<sup>46</sup>.

A propos des synonymes grecs *poti* : *proti* « à côté de », Chantraine affirme dans son dictionnaire qu'ils n'ont pas « la même origine », affirmation qui n'est accompagnée d'aucun argumentaire.

Brugmann est plus prudent :

Il est moins vraisemblable de supposer que \**poti* d'époque indo-européenne soit sorti de \**proti*, que de croire qu'il a été sorti de \**po*, et que dans l'usage il s'est en quelque sorte confondu avec \**proti*, qu'il en a usurpé les autres fonctions<sup>47</sup>.

Dans les faits, ces deux mots sont inséparables, mais dans la doxa, étant donné que -r- n'a pas le statut d'infixe ou d'incrément, la seule posture théorique admissible est de nier par principe qu'ils aient un lien génétique possible. Notre diagnostic est que cette posture n'est pas tenable compte tenu de l'accumulation de données qui montre que l'incrémentation est un phénomène ultra-fréquent en indo-européen.

---

44. L'auteur n'imagine pas qu'il puisse s'agir d'une incrémentation.

45. Le sigle Grdr renvoie à Brugmann et Delbrück, *Grundriss der vergleichenden Grammatik der indo-germanischen Sprachen* [traité de grammaire comparée des langues indo-européennes].

46. Boisacq (1923, p. 8).

47. Brugmann (1905, p. 504, § 612). Cet exemple gêne (notez le mot « usurpé »).

Prenons les synonymes anglais suivants, qui signifient tous « sauter, bondir ».

- to *skip* ME *skippen* origine inconnue
- to *hop* OE *hopp-ia* racine i.e. *keu-p*
- to *leap* OE *leap-a* racine i.e. *klou-p*

Si on fait l'analyse que #s- est le préfixe #s- mobile et que -l- est un infixé, alors ces trois formes supposées indépendantes se ramènent à une racine unique  $\sqrt{*k\_b}$  qui possède trois formes :

- s-k\_b d'où *skip*
- k\_b d'où *hop*
- k-l-\_b d'où *leap*

L'économie dans la description est évidente. En outre, cette analyse intègre *to skip* dans une famille de mots, alors qu'il était isolé et d'origine inconnue.

De même en grec,  $\sqrt{*k\_p}$  :

- kalupt-ô* « couvrir, cacher »
- krupt-ô* « cacher »

De même en grec,  $\sqrt{*H_2\_k}$  :

- alek-sô* « protéger, défendre »
- alkê* « défense, protection »
- arke-ô* « protéger »

De même,  $\sqrt{*p\_k}$  :

Grec	<i>ha-paks</i> < *s_m + *p_k-s	« une fois »
Latin	<i>sim-p-l-eks</i> < *s_m + *p-l_k-s	« simple < en une fois <sup>48</sup> »
Allemand	<i>ein-fach</i> < *oin + *pak	« simple < en une fois »

Prenons le mot *peigne* en slave : les différentes formes reposent sur \*greb- avec divers suffixes. Ces formes slaves sont à première vue isolées dans l'espace indo-européen. Une fois reconnu que -r- est un incrément, on retrouve la racine  $\sqrt{*g\_bh}$ , qui donne les formes germaniques comme l'anglais *comb* « peigne », et vieux norrois *kambr*, avec un autre incrément.

Non seulement les langues indo-européennes sont apparentées entre elles, mais elles le sont à un degré qui est bien plus fort que ce que le comparatisme classique permet de décrire. Du fait d'une théorie de la racine, défailante et inadaptée, quantité de lexèmes sont tenus – à tort – pour isolés ou non indo-européens.

Une autre racine  $\sqrt{*bh\_g}$  *casser, frapper* :

racine i.e. <i>bheg</i>	« asséner un coup »	sanscrit <i>bhanj</i> ; arménien <i>bek-anim</i>
racine i.e. <i>bhlag</i>	« frapper de façon répétée »	latin <i>flagere</i>
racine i.e. <i>bhlîg</i>	« frapper de façon répétée »	latin <i>flîgere, flictus</i>
racine i.e. <i>bhreg</i>	« casser »	latin <i>frangere, fractus</i> (-r- péjoratif)

48. Littéralement, « fait en un seul pli ».

Une autre  $\sqrt{*H_1-p}$  : *saisir, prendre en main* :

racine i.e. Hop « choisir » latin *op-târe*  
 racine i.e. (H)rep « s'accaparer, voler » latin *rap-ax, rap-t* (-r- péjoratif)  
 racine i.e. (H)leup sanscrit *lop-tra* « butin »

Un autre exemple plus complexe  $\sqrt{*bh_H_2}$  *briller, être blanc ou blond* :

racine i.e. bhaH<sub>2</sub> « briller » grec *phôs* « lumière »  
 racine i.e. bhleH<sub>2</sub>-k « briller, blanc » anglais *bleach, bleak*  
 racine i.e. bhleH<sub>2</sub>-nk « briller, blanc » français *blanc*  
 racine i.e. bhreH-k-t « briller » anglais *bright*  
 racine i.e. bhleH<sub>2</sub>-nt « briller, blond » anglais *blond*  
 racine i.e. bhleH<sub>2</sub>-w « briller, blond » latin *flavus* « blond-roux »

Ou  $\sqrt{*k_w}$  *être très sonore, faire du bruit*

i.e. keu- anglais *shout* < ME *shoute* < \*s-kw-d-  
 i.e. k-l-eu-eH<sub>1</sub> anglais *loud* < \*hlûd- < \*kwH<sub>1</sub>-tos  
 i.e. k-r-eu (-H<sub>2</sub>) instrument de musique : gallois *crwth* < \*kruttâ

Plusieurs racines présentent le paradigme incrémental suivant :

Préfixe	Infixe	Suffixe
#s-	-r- -l-	-s-

Par exemple :  $\sqrt{*gh\_bh}$  *tenir en main* :

racine nue anglais *to give* < OE *giba* < \*ghebh  
 racine suffixée latin *hab-êo* < \*ghabh-eH<sub>1</sub>-  
 racine infixée anglais *to grab* < néerlandais *grabben* < \*gh-r-obh  
 racine infixée anglais *to grasp* < \*gh-r-obh-s  
 (avec métathèse normale en anglais)

Par exemple :  $\sqrt{*bh_H}$  *pousser, croître ; arbre*<sup>49</sup> ; *fleur* :

racine suffixée grec *phêgos* « chêne », anglais *beech* « hêtre » < \*bh\_H-g  
 racine suffixée allemand *Baum* « arbre » < \*bh\_H-m  
 racine infixée anglais *birch* « bouleau » < \*bhrH-g<sup>50</sup>  
 racine infixée latin *fraxinus* « frêne » < \*bhrH-g-s-  
 racine infixée allemand *Blüte* « fleur » ; latin *flor* « fleur »

49. Le signifié « arbre » pour la racine nue est attesté en ouralien : cf. finnois *puu* « arbre ».

50. Ce mot a été maintes fois convoqué dans l'épuisante quête de la patrie originelle indo-européenne. Il nous apprend que dans cette patrie, il y avait des arbres. Mais on ne sait pas très bien lesquels.

### L'incrément #d < t' ou ṭ

La théorie classique ne rend pas compte des infixes -r- et -l-. Une autre absence concerne le préfixe #d-<sup>51</sup>. L'exemple canonique est \*dakru- « larme » en face de \*akru « larme ». A ce jour, les comparatistes ne se sont pas résolus à postuler un préfixe #d-. Il existe plusieurs exemples très clairs :

<i>akru</i>	« larme »	<i>d-akru</i>	« larme (exemple canonique) »
<i>rew</i>	« couler »	<i>d-rew</i>	« s'écouler »
<i>ar-bor</i>	« arbre »	<i>d-oru</i>	« arbre »
<i>nebh</i>	« nuage »	<i>d-nephos</i>	« ténèbres » (grec)
<i>ors</i>	« derrière, cul »	<i>d-ors</i>	« dos » (latin)
<i>eigh</i>	« être piquant »	<i>d-eigh</i>	« piquer, être piquant »

A notre avis, ce incrément est représenté en arabe.

Nous en proposons quelques exemples ci-dessous comme préfixe.

Arabe *baḥar* « fumer ; fumée, vapeur d'eau bouillante »<sup>52</sup>

Arabe *ṭabaḥ* « cuire »

Arabe *bašar* « homme, genre humain »

Arabe *ṭabš* « hommes, genre humain »

Arabe *ba'ar* « déféquer, excréments »

Arabe *ṭaba'* « saleté, crasse, boue »

Arabe *baqar* « fendre, ouvrir »

Arabe *ṭabaq* « séparer deux parties à la charnière »

Arabe *rašša* « arroser, asperger »

Arabe *ṭaraša* (F. II) « répandre ça et là (en arrosant, en aspergeant) »

Arabe *radda* « repousser, éloigner, écarter »

Arabe *ṭarada* « éloigner, écarter, repousser, chasser »

D'autre part, il existe au moins un exemple de -ṭ- infixé.

Racine √ H\_m « être en colère »

Arabe *ḥaṭam* (F. V) « être enflammé de colère »

Arabe *muḥṭamir* « courroucé »

Arabe *ḥamâ* « être en colère »

Arabe *ḥamâ'* « être en colère »

Arabe *ḥamal* « montrer en colère ; exciter la discorde et l'inimitié »

Arabe *ḥamaq* (F. V) « se fâcher »

Arabe *ḥamar* « s'enflammer de colère »

Arabe *ḥamas* (F. V) « mettre en colère »

Arabe *ḥamašâ* « irriter »

Arabe *ḥamiya'* « emportement de colère, début de colère »

Arabe *ḥašma'* « colère »

Arabe *ḥadam* (F. IV) « s'enflammer de colère » ; (F. VIII) « bouillonner de colère »

51. Ce préfixe est attesté surtout sur des racines de sens médio-passif, ce qui permet le rapprochement de cet incrément avec le préfixe passif du berbère : #ṭ-.

52. Cf. i.e. *bhok* « feu » > latin *focus* « feu », arménien *bok'* « feu ». Et arabe *nabīḥa'* « allumette ».

## L'incrément -t-

L'incrément -n- n'est pas le seul infixé de l'indo-européen. Il y a de nombreux exemples en grec d'un infixé -t-. Ainsi :

*ptolemaios* « le belliqueux » ~ *polemos* « la guerre »

*ptolis* « la ville » ~ *polis*

*ptaiô* « heurter, renverser » ~ *paiô* « battre »<sup>53</sup>

Depuis le XIX<sup>e</sup> siècle, les comparatistes eux-mêmes ont observé des exemples possibles, mais ils n'ont jamais osé franchir le pas et postuler des incréments. Ici aussi, on arrive à expliquer des termes isolés et à déterminer des racines nouvelles.

Un autre corpus ingérable dans l'approche orthodoxe :

Grec *ptelea* « orme »

Latin *populus* « peuplier »

Russe *topol'* « peuplier »

Examen des lexèmes grecs :

*pterna* « talon » ~ skrt *parśnis* (exemple classique repéré depuis le XIX<sup>e</sup> siècle)

*pteron* « aile » < \**pet* « voler » ~ latin *penna* < \**petna* (ici -t- est radical)

*ptissô* « piler, concasser » ~ latin *pis-ton* ; « pilon » < \**pislum* ; latin *pinsô* « piler » (infixe -n-)

*pto(i)a* « peur, épouvante » ~ latin *pau-ra*

*ptuk<sup>h</sup>s* « pli » ~ allemand *biegen* « plier »

*ptûô* « cracher, vomir » ~ latin *spu-ô* ; anglais *to spew*

*ptôkh-eô* « mendier » ~ anglais *to beg* (ces deux mots sont isolés)

*kti-zô* « bâtir une maison » ~ i.e. *kei-m*, germanique *haim* « maison, home »

A notre avis, l'incrément -t- existe également en sémitique.

Nous en proposons quelques exemples ci-dessous.

Racine  $\sqrt{r\_}$  « paître »

Arabe *ra'â* « paître »

Arabe *rata'* (F. IV) « laisser paître librement »

Racine  $\sqrt{m\_š}$  « traire »

Arabe *mašša* « traire une femelle en lui laissant un peu de lait dans le pis »

Arabe *mataš* « presser doucement les trayons d'une femelle en la trayant »

Racine  $\sqrt{l\_h}$  « sale »

Arabe *latah* « salir quelqu'un d'ordures »

Arabe *lahh* « oindre, imprégner »

Arabe *lahhatun* (fém.) « sale, malpropre et qui sent mauvais (femme) »

Hébreu *lih<sup>h</sup>lûh* « saleté, crasse »

Hébreu *lih<sup>h</sup>leh* « salir, souiller »

Racine  $\sqrt{h\_l}$  « s'écouler, se répandre abondamment (pluie, larmes) »

Arabe *halla* (F. VII) « être versé par torrents, se répandre (pluie) »

Arabe *halla* (F. VII) « être baigné de larmes (yeux) »

---

53. Cf. Meillet : « [...] dans *paiô*, non plus que dans *ptaiô*, qu'on n'en peut séparer [...] », Ernout et Meillet (1932, p. 708).

Arabe *hatal* « faire tomber par intervalles des ondées, des averses (ciel) »  
Arabe *hatal* « verser des larmes abondantes »  
Arabe *hatîl* « qui fait tomber une pluie continuelle (ciel) »

Racine √ h\_m « noir »

Arabe *hutma'* « couleur noire »  
Arabe *hâtim* « corbeau »  
Arabe *ḥamma* « être noir »  
Arabe *'aḥamm* « noir »  
Arabe *'aḥtam* « noir »  
Arabe *ḥamam* « charbon »  
Arabe *ḥamḥam* « très noir »  
Arabe *ḥamâ* (F. XII) « être noir (nuit, nuage) »  
Arabe *yaḥmûm* « noir »  
Arabe *saḥam* « noir »  
Arabe *ḥamâta'* « sang noir dans le cœur »<sup>54</sup>

Racine √ h\_n « couper »

Arabe *ḥanna* « couper (le tronc d'un palmier) »  
Arabe *ḥana'* « couper (le tronc d'un palmier) »  
Arabe *ḥatan* « tronquer, écourter »  
Arabe *ḥanaf* « couper (un fruit rond) »

Racine √ f\_q « ouvrir, fendre en deux »

Arabe *fataq* « fendre, rompre, séparer ce qui était joint par une charnière »  
Arabe *faqqa* « ouvrir ; disjointre, séparer en deux »  
Arabe *faqa'* « fendre ; séparer »  
Arabe *faqaḥ* « ouvrir les yeux pour la première fois (chiot) »  
Arabe *faqa'* (F. VII) « se fendre ; être fendu »  
Arabe *falaq* « fendre, couper en deux »  
Arabe *faraq* « fendre, pourfendre en deux »

Racine √ k\_m « cacher, recouvrir »

Arabe *kamma* « couvrir, recouvrir (d'un couvercle, d'une enveloppe) »  
Arabe *katam* (F. IV) « céler, cacher, dérober à la vue »

### Les formes verbales dérivées en indo-européen

A partir d'une base trilitère, l'arabe tire plusieurs formes dérivées :

Forme I	forme de base	fa'al
Forme II	duplication interne	fa'al
Forme III	allongement	fâ'al
Forme IV	préfixation	'af'al

La forme I est représentée en indo-européen : c'est la forme de citation de chaque racine. La forme II existe aussi, mais c'est toujours la première consonne qui est dupliquée et non la deuxième.

---

54. Ce serait un exemple d'incrémentation suffixale par -ṭ.



Par exemple : √ p\_l « plein, abondant ; foule, multitude »

Latin *populus* « foule, peuple »

Anglais *folk* « peuple »

Arabe *ḥafla* « foule »

Latin *plēnus* « plein » < \*pleH<sub>1</sub>-nos

Grec *pimplē-mi* « je remplis » (-mi : « je, P1 »)

Arabe *ḥafil* « plein »

Pour des raisons qui restent à élucider, la forme II est souvent associée en indo-européen au passé. Par exemple, en latin : *mordēo* « je mords », *momordi* « j'ai mordu » ; *dō* « je donne », *dedi* « j'ai donné ».

La forme III à allongement est bien représentée. Il s'agit des degrés vocaliques dits longs /e:/ /o:/, dont l'origine est l'allongement de la voyelle simple /e/ /o/ et non la trace d'une laryngale /e+H/ ou /o+H/. La classe dite VII des verbes forts germaniques repose sur la forme III. Le verbe anglais *to bear* « porter » appartenait à cette classe : *I bear* < \*bêr « je porte » et *I bore* < \*bôr « j'ai porté ». L'adjectif *wet* « mouillé » repose sur \*/wêd/ avec une longue.

La forme IV est très intéressante. Son existence en indo-européen est niée. Les exemples sont tenus pour des formations aberrantes, souvent qualifiées de « populaires »<sup>55</sup> et dépourvues de toute portée vis-à-vis de la théorie de l'indo-européen. Une autre façon de refouler les exemples en dehors du périmètre indo-européen consiste à en faire des emprunts à un substrat. Il en existe d'excellents exemples, en particulier en grec, mais pas seulement. Cette formation fournit beaucoup de noms d'animaux ou d'objets. La cohérence morphologique et sémantique de cette forme permet d'exclure le hasard comme principe d'explication. En outre, il est fréquent que la racine de ces mots soit au degré zéro, ce qui laisse à penser que le préfixe était autrefois accentué. Nous citons en note le point de vue de Chantraine dans son *Dictionnaire étymologique de la langue grecque*. Le terme convenu pour désigner l'incrément #a- est celui de « prothèse », dont l'origine médicale montre à quel point les indo-européanistes considèrent cet incrément comme un *corps étranger*.

Grec *a-glīs*<sup>56</sup> et *gelgis* « tête d'ail » < \*√ g\_l\_y (exemple d'alternance gel ~ ?a-g\_l)

Grec *a-glaFos* et *glaukos* « brillant » < \*√ g\_l\_w

Grec *a-greiphna*<sup>57</sup> « herse, rateau » et *grīphos* « filet » < \*√ gh\_bh « saisir, attraper »

Grec *aielouros* < \*a-wisēl-<sup>58</sup> « chat sauvage », latin *vison* « vison », anglais *weasel* « belette » < \*√ wis-

Grec *a-i-gupios*<sup>59</sup> et *gups* « vautour »

Grec *a-mal-os*, *a-blēkhros* et *blēkhros* « faible, doux » < \*√ m\_l

Grec *a-nepsios*<sup>60</sup> « cousin germain » < \*√ nep-ot « neveu »

---

55. Sous la plume des comparatistes, en particulier de Meillet, le label « populaire » équivaut à être rejeté dans les ténèbres de la linguistique.

56. Cf. Chantraine : « ne peut être séparé de *gelgis* (voir ce mot) ».

57. *Idem* : « l'α- étant une prothèse non autrement expliquée ».

58. L'élément final *ouros* signifie « queue ».

59. Le -i- après #a- serait dû à l'influence d'autres noms d'oiseaux commençant par #ai- (Chantraine).

Grec *a-pion*<sup>61</sup> < \*a-pis-on « poire » et latin *pirus* < \*pisos « poire »  
 Grec *a-sp-is* « vipère aspic » et latin *serpens* « serpent » < \*√ s\_p « ramper, serpent »  
 Grec *a-spalaks* et *spalaks* « taupe »  
 Grec *a-s-pharangos*<sup>62</sup> « gosier, gorge » et *pharungs* « gosier » < \*√ bh\_r-  
 Grec *a-staphis*<sup>63</sup> et *staphis* « raisin sec », *staphulê* « grappe de raisin »  
 Grec *a-stakhus* et *stakhus* « épi de blé »  
 Grec *a-stralos*<sup>64</sup> et latin *sturnus* « étourneau » < \*√ st\_r-/tr\_s- « étourneau, grive »  
 Grec *a-tharê* « bouillie de farine ou de gruau » < \*ghrew « gruau »<sup>65</sup>  
 Grec *a-nthrênê* « bourdon », *tenthrene* « guêpe », *thrônaks* « bourdon » < \*√ dhren  
 « bourdonner »  
 Grec *a-trapos*<sup>66</sup> « sentier » et *trapeô* « fouler, marcher sur »

A noter qu'en combinant \*?a- et \*-t-, on donne une cohérence √ k\_H aux formes suivantes :

Latin *a-qui-la* « aigle »  
 Grec *a-isa-los* < \*?a-k-ya-los « faucon émerillon »  
 Grec *ik-t-înos* « milan » (avec infixes -t)  
 Germanique commun \*kû-tya « milan »

La variante \*?i- du préfixe \*?a- se retrouve dans d'autres mots :

Irlandais *áth* « four » < Celte \*apatinos < \*?a- + kwH<sub>2</sub>- tin-os, d'après \*√ kwH<sub>2</sub>  
 « cuire, brûler »  
 Grec *ipnos* « four » < \*?i- + kwH<sub>2</sub>- n-os (sans -ti-)  
 Grec *i-erâks*, *îrêks* « faucon » < \*√ H<sub>1</sub>\_r « aigle, grand oiseau »  
 Grec *i-gn-uê* « pli du genou » < \*√ g\_n(H<sub>1</sub>) « engendrer »<sup>67</sup>  
 Grec *gonu* et *gnuks* « genou »  
 Grec (dialectal) *i-kn-us* « cendre » < \*√ k\_n « cendre »<sup>68</sup>  
 Grec *i-khthu:s* « poisson » < \*√ gh-dh-uH « poisson » < \*√ gh\_H « ouvrir la bouche,  
 être béant »<sup>69</sup>  
 Grec *i-skh-ion* « hanche, col du fémur » < \*i-s\_k-snos<sup>70</sup>

60. Cf. Chantraine : « L'α- initial présente l'ambiguïté habituelle, mais semble devoir être interprété comme une prothèse (ou un ə<sub>2</sub>). »

61. *Idem* : « Il faut admettre un thème \*piso, l'α initial fait difficulté comme souvent (prothèse ?). »

62. *Id.* : « Le sens précis du mot conduirait à le rapprocher de *pharungs*. »

63. *Id.* : « Le thème fait penser à celui de *staphulê* "grappe de raisin". La forme à initiale α- semble la plus ancienne. Est-ce une prothèse ? »

64. *Id.* : « Avec prothèse, se rapproche évidemment de v.h.a. *stara*, lat. *sturnus*. »

65. En admettant une métathèse \*ghrew > \*ghwer > θer.

66. Cf. Chantraine : « Il faut admettre un α- copulatif et la racine qui se trouve dans *trapeô* "fouler" [...] c'est la piste foulée. »

67. Ce mot *ignus* est généralement considéré comme provenant de \*en+g\_n- par préfixation. Néanmoins, cette hypothèse pose des problèmes phonétiques car on attend \*\**innus* plutôt que *ignus*.

68. La forme normale en grec est *konia*.

69. Ce mot grec est extraordinaire sur le plan morphologique : incrément \*-dh- et schème vocalique \*i\_u. Les autres attestations n'ont pas l'incrément mais ont la voyelle \*u.

70. Ce mot a été comparé depuis longtemps avec le mot sanscrit *sak-thi* « cuisse », mais la théorie orthodoxe ne permet pas de rapprocher ces deux mots de façon satisfaisante.

A comparer également les mots suivants :

Grec *i-orkos* « chevreuil » (incrément #i-)

Grec *d-orkas* « chevreuil » (incrément #d-)

Anglais *roe* < germanique \**roik-os* « chevreuil » (infixe -i-)

Celte \**y-ork-os* « chevreuil » (incrément #i-)

Ces mots sont sans doute en rapport avec la racine \* $\sqrt{H_1_r}$  « animal à corne », qui donne également le mot latin *ariês* « bélier ».

L'existence de cet incrément en germanique, en italique et en celte, à côté du grec, permet de reporter ce procédé morphologique au stade le plus ancien de l'indo-européen. On ne peut pas dire avec une certitude définitive si #a- \* $H_2e$  remonte à \* $?a-$ , en raison de la multiplicité des valeurs phonétiques de  $H_2$ , mais c'est *a priori* l'hypothèse la plus simple et la plus fructueuse. C'est celle que nous retenons.

Peter Schrijver s'est intéressé à plusieurs lexèmes désignant entre autres des oiseaux dans les langues d'Europe de l'Ouest. Ce corpus fait apparaître l'existence d'un préfixe #a-. L'explication usuelle selon laquelle on toucherait là un substrat non indo-européen nous paraît infondée et surtout inutile. Notre avis est que ces lexèmes sont pleinement indo-européens et constituent des exemples supplémentaires de la forme IV. Une fois de plus, une théorie adaptée de la morphologie fossile permet de donner du sens, d'organiser les données et de simplifier les hypothèses, en bref de rationaliser.

\**mes\_l-* ~ \**a-m\_sl-* « merle »

Gallois *mwyalch*

Latin *merula* > français « merle »

Vieux haut allemand *amsla, amasla, amisla, amusla*

Vieil anglais *o:sl*

\**laHw-* ~ \**a-laHw-* « alouette »

Vieil anglais *la:verce* > *lark*

Vieux haut allemand *le:rahha, le:rihha*

Moyen néerlandais *leewerke*

Latin (d'origine gauloise) *alauda*

Emprunt en finnois (ouralien) *leivo(nen)*

\**raud-* ~ \**a-rud-* « (morceau de) minerai »

Latin *raudus* "lump of ore"

Vieux haut allemand *aruz, ariz*

Vieil anglais *arut*

Emprunt en finnois (ouralien) *rauta*

Emprunt en lapon (ouralien) *ruowde*

## Conclusions et perspectives

La famille indo-européenne est une invention autonome remontant à la Renaissance.

Par la suite, tous les progrès significatifs dans l'étude et la théorisation de cette famille ont été accomplis chaque fois que la modélisation linguistique de l'indo-européen s'est rapprochée de la réalité du sémitique et de l'arabe. Le système phonologique avec des *laryngales* imaginé par Saussure, puis développé par

Möller et Cuny, et la théorie de la racine de Benveniste sont formés d'après le modèle du sémitique, de façon quasiment explicite.

Le proto-lexique de l'indo-européen, reconstruit par les comparatistes, est atomisé en une myriade de racines de forme et de sens proche. De multiples exemples montrent que le lexique présente une organisation propre que la doxa officielle n'a pas su mettre en évidence. La seule façon de sortir de cette impasse est d'appliquer une méthodologie d'analyse identique à celle que propose Georges Bohas pour l'arabe : déterminer des racines et des incréments. Nos analyses montrent que les incréments valides pour l'indo-européen sont les mêmes que ceux qu'on détermine en sémitique, à commencer par -r- et -l-, et aussi -n- ou #?a-. A l'inverse, on trouve en arabe des incréments nouveaux, -t- et -ṭ-, que la TME n'a pas proposés mais dont le corpus indo-européen offre des exemples limpides. La recherche en indo-européen de l'équivalent des formes verbales dérivées de l'arabe permet de donner à la prétendue *prothèse* vocalique #a- un vrai statut linguistique, puisqu'elle correspond à la forme IV.

Par rapport à la question de l'appareillement de l'indo-européen vis-à-vis de l'ouralien ou du chamito-sémitique, nous sommes convaincu que le tropisme sémitisant de l'indo-européen contient implicitement la réponse. Nous avons examiné de près l'ouralien et on cherchera en vain des préfixes et des infixes -r- et -l- dans cette famille. L'approche *Racines et incréments* est également valide et féconde dans le domaine ouralien, mais elle ne fournit pas les mêmes résultats qu'en indo-européen et qu'en sémitique, en termes de morphologie fossile. Ce n'était pas l'objet que nous voulions aborder ici.

Il faut noter qu'il existe, dans le domaine comparatiste, des résistances considérables à certaines évidences. Il a fallu un siècle pour que les Allemands tiennent compte des études de Saussure remontant à 1870. Seebold, dans les années soixante-dix, est le premier à pratiquer le formalisme laryngaliste, trois générations après Saussure, deux générations après la découverte du hittite et une génération après Benveniste. Formulons le vœu que cette fois-ci l'appropriation de l'approche *Racines et incréments* dans le domaine indo-européen soit plus rapide.

## Bibliographie

- BEEKES Robert, 1995, *Comparative Indo-European Linguistics. An Introduction*, Amsterdam, John Benjamins.
- BENVENISTE Emile, 1935, *Origines de la formation des noms en indo-européen*, Paris, Maisonneuve.
- BOHAS Georges, 1997, *Matrices, étymons, racines*, Leuven, Peeters.
- BOISACQ Emile, 1923, *Dictionnaire étymologique de la langue grecque*, Paris, Klincksieck.
- BRUGMANN Karl, 1905, *Abrégé de grammaire comparée des langues indo-européennes*, Paris, Klincksieck.
- CHANTRAINE Michel, 1983, *Dictionnaire étymologique de la langue grecque*, Paris, Klincksieck.
- COLLINGE Neville E., 1985, *The Laws of Indo-European*, Amsterdam, John Benjamins.

- ERNOULT et MEILLET, 1932, *Dictionnaire étymologique de la langue latine*, Paris, Klincksieck.
- GRIMM Jakob, 1848, *Geschichte der Deutschen Sprache*, Leipzig, Herzl.
- KIMBALL Sara E., 1999, *Hittite Historical Phonology*, Innsbrucker Beiträge zur Sprachwissenschaft, n° 95, Innsbruck.
- KURYŁOWICZ Jerzy, 1956, *L'apophonie en indo-européen*, Wrocław, Ossolineum.
- LEHMANN Winfred P., 1955, *Proto-Indo-European Phonology*, Austin, Univ. of Texas Press.
- LEIBNIZ G. W., 1990, *Nouveaux essais sur l'entendement humain*, Paris, Flammarion.
- MALLORY James, 1997a, *A la recherche des Indo-Européens*, Paris, Le Seuil.
- , 1997b, *Encyclopedia of Indo-European Culture*, Fitzroy Dearborn Publishers.
- MARTINET André, 1986, *Des steppes aux océans*, Paris, Payot.
- MEIER-BRUEGGER Michael, 2002, *Indogermanische Sprachwissenschaft*, Berlin, Walter de Gruyter.
- MELCHERT H. Craig, 1994, *Anatolian Historical Phonology*, Leiden Studies in Indo-European.
- MÖLLER Hermann, 1906, *Semitisch und Indogermanisch*, Copenhagen, H. Hagerup.
- MUKHERJEE S. N., 1968, *Sir William Jones*, Cambridge University Press.
- PARSONS James, 1767, *The Remains of Japhet, Being Historical Enquiries into the Affinity and Origins of the European Languages*, Londres.
- POKORNY Julius, 1959, *Indo-Germanisches etymologisches Wörterbuch (IEW)*, Berne, Francke Verlag.
- PUHVEL Jaan, 1984, *Hittite Etymological Dictionary*, Berlin, Mouton Publishers.
- SALMONS Joseph C., 1993, *The Glottalic Theory. Survey and Synthesis*, *Journal of Indo-European Studies*, Monograph Series, n° 10, McLean, Institute for the Study of Man.
- SCHRIJVER Peter, 1991, *The Reflexes of the Proto-Indo-European Laryngeals in Latin*, Amsterdam, Rodopi.
- SAUSSURE Ferdinand de, 1870, *Mémoire sur le système primitif des voyelles dans les langues indo-européennes*, Leipzig, Teubner.
- SCHLEGEL Friedrich, 1808, *Ueber die Sprache und die Weisheit der Indier*, Heidelberg, Mohr und Zimmer.
- SERGENT Bernard, 1995, *Les Indo-Européens*, Paris, Payot.
- SZEMERENYI Oswald, 1973, *La théorie des laryngales de Saussure à Kuryłowicz et à Benveniste*, BSL, n° 68.