



HAL
open science

Linguistique du genre

Jean-Philippe Watbled

► **To cite this version:**

Jean-Philippe Watbled. Linguistique du genre. Geoffroy S., Féral C., Jorrand S. et Bosquet M.-F. Genre et dynamiques interculturelles: la transmission, L'Harmattan, pp.167-179, 2012, 978-2-336-00373-3. hal-00905489

HAL Id: hal-00905489

<https://hal.science/hal-00905489>

Submitted on 27 Sep 2017

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

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

Jean-Philippe Watbled
Université de La Réunion

Linguistique du genre

Le langage naturel nous permet de classer les êtres qui nous entourent, y compris nos semblables. L'un des paramètres de cette classification est le genre qui, se trouvant à la croisée des chemins de la nature et de la culture, est l'objet de représentations mentales et idéologiques, ce qui ne manque pas de susciter des débats, plus ou moins vifs selon les époques. J'aborde ici cette question principalement sous l'angle de la linguistique et du comparatisme, avec l'idée que les principes lexicaux, morpho-syntaxiques et sémantiques à l'œuvre dans les systèmes de classification sont les prolégomènes à toute étude plus vaste du genre. Nous devons nous demander si les catégorisations s'appliquent aux noms ou bien aux entités auxquelles ils réfèrent, ou encore aux deux. Notons aussi dès à présent que, dans le cadre de l'étude proprement linguistique du genre, envisagée ici comme prioritaire, se pose le problème classique de l'opposition entre relativisme et universalisme.

Les systèmes linguistiques

Toute représentation passe par le langage, trait distinctif de l'humain¹, qui sert d'interface entre le sujet et le monde que ce sujet perçoit et cherche à connaître. Cependant, malgré cela, les mots et les choses sont hétérogènes. En outre, la syntaxe sert elle-même d'interface, interne au langage cette fois, entre le plan de l'expression signifiante et le plan du contenu signifié. À cet égard, les langues oscillent entre iconicité et arbitraire structural².

Le linguistique, on s'en doute, est en relation avec l'extralinguistique, car il s'agit bien, avant tout, de parler de quelque chose d'extérieur au langage, mais il est évident qu'il n'existe pas d'isomorphie stricte entre d'une part les catégories propres à une langue donnée, et d'autre part d'hypothétiques catégories primitives, universelles ou supposées telles, impliquées dans l'analyse du monde. D'ailleurs, si cette isomorphie existait systématiquement, l'espèce humaine ne parlerait en quelque sorte qu'une seule langue sur cette planète, ou du moins sous des langues apparemment étrangères les unes aux autres se

¹ Voir par exemple la place que Descartes attribue à la parole (*Discours de la méthode, Méditations métaphysiques*).

² Pour le langage naturel, on parle d'iconicité quand il existe une relation analogique entre le plan de l'expression et le plan du contenu. L'arbitraire s'oppose à l'iconicité.

cacheraient une grammaire unique, et les disparités entre les idiomes ne consisteraient qu'en des différences de couplage entre une grammaire sous-jacente universelle et des lexiques nécessairement diversifiés en raison de l'arbitraire du signe³.

Deux points de vue

La complexité des langues a pour conséquence que le genre comme catégorie linguistique doit s'envisager de deux points de vue distincts : (i) le genre sémantique ; (ii) le genre lexico-grammatical. Il importe en outre de distinguer le genre (sémantique ou lexico-grammatical) et la manière dont il se manifeste dans l'expression linguistique.

Le point de vue sémantique s'appuie largement sur les données biologiques mais, alors que le genre naturel se limite (en principe) à deux possibilités (mâle, femelle), les langues peuvent aussi faire intervenir d'autres propriétés, comme animé/inanimé, par exemple, couple qui n'est habituellement pas considéré comme relevant du genre naturel : celui-ci peut donc ne constituer qu'un sous-ensemble du genre sémantique. En outre, on peut faire référence à un animal ou à un humain indépendamment de son genre naturel, en ne prenant en compte que l'espèce à laquelle il appartient. Par ailleurs, telle espèce animale pourra être « masculinisée » ou « féminisée », quel que soit le genre biologique des individus. Enfin, on se gardera d'oublier que la conception sémantique du genre, qu'il s'agisse de mâle/femelle ou d'animé/inanimé, par exemple, est susceptible de différer d'une culture à l'autre⁴.

Le genre lexico-grammatical se caractérise par l'existence de classes relevant à la fois du lexique et de la grammaire, et qui se manifestent souvent par des mécanismes tels que l'accord (cf. *le rat, la souris*), ou par la forme des pronoms personnels⁵ (cf. *il* et *elle* en français), mais qui, selon les langues, ne sont pas purement sémantiques, ou ne le sont pas du tout.

Pour comprendre la différence entre le genre sémantique et le genre lexico-grammatical, le plus simple est de comparer le français et l'anglais. En français, même s'il existe une concordance certaine entre genre sémantique et genre lexico-grammatical dans la référence aux humains (voir plus loin), les pronoms *il* et *elle* renvoient dans la plupart des cas au genre

³ Cette notion est due à Ferdinand de Saussure, *Cours de linguistique générale*, Paris, Payot, 1980 (initialement publié en 1916). Dans l'hypothèse dont il est ici question, cela signifierait que l'arbitraire ne concernerait que les unités lexicales, mais pas la syntaxe. En fait, ma position personnelle est qu'il existe des catégories primitives et des principes universels, mais que les unes et les autres sont organisés différemment dans chaque langue.

⁴ Nous reviendrons plus loin sur ces deux derniers points à propos de l'indo-européen.

⁵ Rien n'empêche que ces procédés manifestent aussi le genre sémantique. Comme le genre lexico-grammatical n'a pas de fondement purement sémantique, il faut bien qu'il se manifeste par des procédés formels, sous peine d'être indécélable. En bref, les manifestations formelles du genre sémantique ne sont pas une nécessité *a priori*, puisque le plan notionnel peut suffire, alors que celles du genre lexico-grammatical le sont.

lexico-grammatical (cf. *il* pour le bureau, *elle* pour la table), tandis qu'en anglais, les pronoms *he*, *she* et *it* renvoient directement au genre sémantique : *he* pour une personne de sexe masculin, *she* pour une personne de sexe féminin, et *it* dans les autres cas⁶. Les formes des pronoms en anglais constituent donc une manifestation du genre sémantique⁷. En revanche, en français, les formes de ces pronoms sont en harmonie avec les règles d'accord en genre (lexico-grammatical), inexistantes en anglais⁸ : cf. *le bureau* → *il*, *la table* → *elle*.

Motivation, concordance, discordance

Lorsqu'une langue ne connaît que le genre sémantique, comme l'anglais, ce genre est nécessairement motivé par nature. En revanche, le genre lexico-grammatical est forcément au moins en partie immotivé, ce qui signifie que l'on ne peut postuler un tel système que si les classes (telles que masculin, féminin, neutre) ne correspondent pas systématiquement au genre sémantique⁹. À ce propos, il convient d'observer que dans des langues comme le français ou l'allemand, les noms désignant les entités asexuées peuvent parfaitement être de genre lexico-grammatical masculin ou féminin. Dans ce type de système, on constate par définition des écarts entre, d'une part, le genre sémantique et, d'autre part, le genre lexico-grammatical. Je citerai comme exemple *Mädchen* en allemand : ce nom est neutre à cause du suffixe (*-chen*), mais il reste bien entendu féminin du point de vue sémantique. Ce problème est inconnu de l'anglais, qui ne connaît que le genre sémantique.

En français, les discordances entre les deux plans sont particulièrement patentes pour les noms désignant des êtres sexués. Il existe certes de nombreux cas de concordance, c'est-à-dire de genre lexico-grammatical motivé, tels que *un homme*, *une femme*, ou *un tigre*, *une tigresse* : dans le premier couple, seul l'accord indique le genre, et masculin et féminin correspondent ici à mâle et femelle¹⁰ ; le principe est le même pour le second couple, mais il s'y ajoute une manifestation suffixale du genre¹¹. Il arrive aussi qu'un même lexème connaisse

⁶ Certes, en français, *il* et *elle* peuvent renvoyer apparemment au genre sémantique : *un homme* → *il*, *une femme* → *elle*, mais en fait c'est le genre lexico-grammatical qui reste ici crucial. Il existe des cas problématiques, comme *la sentinelle* pour désigner un garde de sexe masculin : dans une reprise, emploiera-t-on l'anaphore *elle*, en harmonie avec le genre lexico-grammatical, ou l'anaphore *il*, en harmonie avec le genre sémantique ?

⁷ Remarquons à ce propos qu'il existe des langues à genre purement sémantique, mais sans expression pronominale de ce genre : le quechua, par exemple, a un pronom de 3^{ème} personne unique, *pay* (*paykuna* au pluriel), quel que soit le genre sémantique (masculin ou féminin).

⁸ Je précise que rien n'empêche d'avoir un système de genre purement sémantique avec des règles d'accord.

⁹ Le caractère motivé ou immotivé est souvent affaire de degré : le masculin lexico-grammatical de *bureau* n'est pas motivé, mais celui de *matou* l'est.

¹⁰ En réalité, le terme masculin sert aussi à désigner soit un individu indépendamment de son genre naturel, soit l'espèce.

¹¹ Même en français, le genre sémantique peut donc recevoir une expression morphologique, plus précisément morpholexicale (dérivationale en l'occurrence). De même en anglais, on a *tiger*, *tigress*.

deux genres, concordants l'un et l'autre : *un élève, une élève, un enfant, une enfant*. Le changement de classe lexico-grammaticale concorde ici avec le changement de genre sémantique. Mais la discordance entre les deux plans est fréquente : ainsi, *un léopard* peut désigner un mâle ou une femelle, tout comme *une panthère*, alors que sur le plan lexico-syntaxique, le premier est masculin et le second féminin¹².

Classes nominales

Selon les langues, le genre lexico-grammatical se limite ou non au masculin et au féminin : les langues classiques (sanskrit, grec, latin¹³), par exemple, avaient une catégorisation à trois termes de ce point de vue (masculin, féminin, neutre), tout comme l'allemand moderne.

Certaines familles de langues vont plus loin, avec des systèmes de classes nominales impliquant des ensembles de catégories autres ou plus larges que ce que les occidentaux conçoivent comme le genre. Le principe des spécificatifs en chinois en est un exemple. L'emploi d'un spécificatif (ou classificateur) y est obligatoire lorsqu'un numéral ou un démonstratif détermine un nom, et la sélection du spécificatif dépend de la classe du nom. Ce système de classification, qui dépend davantage de la forme des entités que de leur nature, peut être assimilé à un système de genre lexico-grammatical partiellement motivé, mais non lié au genre naturel¹⁴.

Dans un tout autre domaine géographique, les langues bantoues, par exemple, vont également bien au-delà des catégorisations limitées que l'on trouve dans les cas les plus courants. Les noms s'y rangent dans différentes classes lexico-grammaticales¹⁵, qui se reconnaissent par leur préfixe de nombre (singulier ou pluriel), sachant qu'une classe donnée est soit singulier, soit pluriel, certaines fonctionnant par paires. On a ainsi en swahili : *m-tu* (« homme, être humain »), *wa-tu* (« hommes, êtres humains »), *ki-tu* (« chose »), *vi-tu* (« choses »), etc., et en zoulou : *umu-ntu* (« homme, être humain »), *aba-ntu* (« hommes, êtres humains »), *u-baba* (« père »), *o-baba* (« pères »), *isi-kole* (« école »), *izi-kole* (« écoles »), *um-fula* (« rivière »), *imi-fula* (« rivières »), etc. Le swahili a quinze classes nominales et le

¹² Lorsqu'un même lexème peut désigner aussi bien le masculin que le féminin du point de vue sémantique, il est dit épïcène: *panthère, léopard*, mais aussi *élève, enfant*, par exemple, sont épïcènes. Certains considèrent aussi comme épïcènes les emplois génériques du masculin, même quand le féminin existe : *le rat*, par exemple.

¹³ Les latinistes connaissent bien les modèles *dominus*, masculin, *ancilla*, féminin, *templum*, neutre.

¹⁴ Pour le mandarin, voir Jerry Norman, *Chinese*, Cambridge University Press, 1988, p. 115-117 ; Roger Darrobers & Xiaohong Xiao Planes, *Éléments fondamentaux de la phrase chinoise*, Paris, éd. You-Feng, 1998 ; Philippe Roche, *Grammaire active du chinois*, Paris, Larousse, 2007, ch. 9. Pour le cantonnais, on peut consulter Stephen Matthews & Virginia Yip, *Cantonese Grammar*, London & New York, Routledge, 1994, p. 92-107.

¹⁵ Voir Benji Wald, "Swahili and the Bantu Languages", in Bernard Comrie (ed.), *The World's Major Languages*, London & Sydney, Croom Helm, 1987, p. 991-1014.

zoulou dix-huit¹⁶. Précisons que la motivation sémantique n'est que partielle. En outre, le genre naturel ne constitue pas un paramètre donnant lieu à une classe propre. Ainsi, en zoulou le nom *indoda* (« homme », au sens mâle), par exemple, est de la même classe que *intombazana* (« fille »).

Le genre en indo-européen

Pour les mentalités occidentales modernes, le genre sémantique se présente en principe de manière nettement dichotomique : il correspond au genre naturel et présente deux termes, le féminin et le masculin. Or il semblerait bien que cette représentation ne se soit pas toujours imposée aussi nettement. Tout laisse en effet supposer que le genre en proto-indo-européen, ancêtre de la majorité des langues européennes et des langues de l'Inde du nord, entre autres, obéissait à un paramètre différent à la base, qui était le caractère animé ou non des entités telles qu'elles étaient conçues par la communauté.

Il est important de se souvenir que le terme *animé* vient du latin *anima*, le « souffle ». Étaient animés d'un souffle les êtres vivants : humains, animaux, mais aussi les végétaux, contrairement aux minéraux. En outre, les forces naturelles (astres, Terre, fleuves, par exemple) étaient divinisées ou associées à du divin, et par conséquent classées dans le genre animé. La première dichotomie indo-européenne reposait donc sur l'opposition binaire animé ~ inanimé, qui étaient les deux genres fondamentaux. Il s'agissait d'un genre sémantique, mais dans un cadre culturel considéré aujourd'hui comme archaïque par les locuteurs des langues indo-européennes modernes.

L'opposition masculin ~ féminin existait certes en proto-indo-européen, mais était sentie comme seconde par rapport à l'opposition plus fondamentale animé ~ inanimé¹⁷. À date très ancienne, les indices formels de la différenciation masculin ~ féminin au sein de la catégorie des animés restaient faibles¹⁸. En outre, comme en témoignent les langues classiques, filles du proto-indo-européen d'après la linguistique historique orthodoxe, cette opposition est loin de coïncider avec la différenciation sexuelle telle que nous l'entendons aujourd'hui. Ainsi, en latin *terra*, par exemple, est féminin parce que la terre est fécondée.

¹⁶ Voir Clement Martin Doke *et al.*, *English-Zulu, Zulu-English Dictionary*, Johannesburg, Witwatersrand University Press, 1990, p. xvii de la partie *Zulu-English*.

¹⁷ Pour une excellente présentation de cette question, voir l'ouvrage classique de Jean Humbert, *Syntaxe grecque*, Paris, Klincksieck, 1945, 1960, 1972, 2004, p. 9-15.

¹⁸ On en constate des traces dans les langues classiques : on peut songer ici à *pater*, *mater*, par exemple, qui ont des paradigmes flexionnels analogues en latin, bien qu'étant de genres différents.

Dans le passage du proto-indo-européen aux langues classiques, la démotivation s'installe très vite et assez facilement, tant l'opposition masculin ~ féminin sera métaphorisée et vue avant tout de manière imagée. Les noms d'animaux dangereux, par exemple, seront masculins, indépendamment du genre naturel de l'animal, tandis que les noms d'animaux réputés sans bravoure seront féminins. Ainsi, on peut citer le loup (latin *lupus*, masculin), animal virilisé, et le renard (latin *vulpes*, féminin), animal féminisé¹⁹. On passera ainsi insensiblement d'un genre purement conceptuel à une classification lexico-grammaticale en partie immotivée. On est en droit de penser que les métaphores ont été de moins en moins senties comme telles au fil du temps et que, dans le cas du genre, la mort d'une métaphore entraîne forcément un système lexico-grammatical en partie démotivé.

Il subsiste néanmoins une forte affinité entre le genre lexico-grammatical et le genre sémantique, comme l'indiquent des noms tels que *dominus* (« maître ») et *domina* (« maîtresse »)²⁰ en latin. Noms féminins et masculins restent traités à égalité face au neutre, en ayant des formes distinctes pour le nominatif et l'accusatif, quand le neutre connaît le syncrétisme de ces formes²¹. Cette singularité du genre neutre s'explique par les propriétés sémantiques de l'ancien inanimé qui est à son origine. Par nature, il ne possédait pas la puissance requise pour être agent des verbes transitifs dans la phrase d'action, ce qui rendait non nécessaire pour lui une forme distincte de nominatif. Une forme indifférenciée suffisait donc pour l'inanimé sujet des structures intransitives et objet des structures transitives, tandis que masculin et féminin avaient en commun la puissance agentive²², rendant nécessaire cette distinction formelle entre le nominatif et l'accusatif. En effet, dans la phrase d'action à la voix dite active, ces deux cas étaient les expressions morphosyntaxiques des rôles d'agent et de patient, respectivement.

¹⁹ Voir Humbert, *op. cit.*, p. 11.

²⁰ On sait que l'appartenance à un certain type flexionnel n'est pas forcément un indice certain de l'appartenance à une classe morphosyntaxique (= genre) : ainsi *nauta* (« matelot ») est masculin, malgré son type flexionnel largement caractéristique de noms féminins : seul l'accord est révélateur du genre (*bonus nauta*).

²¹ Comme tous les neutres, le nom latin *templum* (« temple »), par exemple, a une même forme pour le nominatif et l'accusatif (*templum* au singulier, *templa* au pluriel). On parle de syncrétisme lorsqu'on a une même forme pour deux propriétés grammaticales distinctes. Seule la confrontation des paradigmes dans un même contexte autorise à postuler le syncrétisme (exemple : le nominatif sujet *dominus* est à l'accusatif objet *dominum* ce que le nominatif sujet *templum* est à l'accusatif objet *templum*).

²² C'est ce fait qui a incité certains linguistes à envisager l'hypothèse ergative pour le proto-indo-européen. On a un système ergatif quand le même cas indifférencié est employé pour le sujet des structures intransitives et pour l'objet direct des structures transitives. Un système morphosyntaxique non ergatif tel que celui du latin ou du grec est appelé système accusatif. Sur l'hypothèse ergative pour l'indo-européen archaïque, voir (entre autres) André Martinet, *Évolution des langues et reconstruction*, Paris, P.U.F., 1975, ch. VIII, et Jean Haudry, *L'indo-européen*, Paris, P.U.F., 1979, p. 97-99.

Comparatisme, relativisme, universalisme

À ce stade de notre étude, la question qui se pose est celle de la relation entre les types de système linguistique et les familles historiques de langues.

On sait que le proto-indo-européen (vu ci-dessus), langue reconstruite et non attestée, mais dont l'existence ne fait aucun doute, était parlé par des peuplades qui ont migré des steppes d'Asie à l'Atlantique²³. Le résultat est que les langues indo-européennes modernes ont hérité d'un même proto-système, ce qui les rend peu différentes les unes des autres pour ce qui est du genre, malgré leur dissémination sur de très larges aires.

Non seulement toutes ces langues de la même famille historique ont un fonctionnement analogue, mais les langues des autres familles historiques des mêmes régions ont également un fonctionnement très proche, avec une opposition très courante entre masculin et féminin. On pourrait citer ici, à titre d'exemples, la famille sémitique, avec l'arabe, l'hébreu ou l'araméen (et d'autres), ou la famille dravidienne avec le tamoul, par exemple.

Prenons le cas de l'arabe du Caire²⁴. Dans ce parler, on a par exemple, *malik* (forme de base) pour « roi » et *malika* (forme suffixée en *-a*) pour « reine » ; le principe est le même pour *kalb* (« chien ») et *kalba* (« chienne »). Les adjectifs connaissent le genre : *kibiir*, *kibiira* (« grand », masc. et fém.). Les exemples cités pourraient à tort laisser entendre que l'on a affaire à un genre sémantique. En fait, il s'agit d'un genre lexico-grammatical largement immotivé, comme l'attestent des noms tels que *maktab* (« bureau », masc.) et *maktaba* (« bibliothèque », fém.). On trouve des noms féminins sans indice de genre : *bint* (« fille »), de genre naturel féminin, mais aussi *shams* (« soleil »), *suu'* (« marché »), dont le féminin est purement lexico-grammatical, puisque ces termes ne désignent pas des animés. On rencontre aussi des cas de supplétion, exactement comme *homme*, *femme*, en français : *raagil* (« homme »), *mara* (« femme »).

Les langues dravidiennes²⁵ opposent l'animé à l'inanimé, ou l'humain au non humain, et au sein de l'animé ou de l'humain, elles opposent le masculin au féminin. Le tamoul oppose un genre « rationnel » à un genre « non rationnel », ces catégories correspondant approximativement à humain ~ non humain. Au sein de la catégorie du « rationnel », on oppose masculin et féminin sur une base strictement naturelle et conceptuelle. Ce système

²³ Voir André Martinet, *Des steppes aux océans*, Paris, Payot & Rivages, 1986, 1994.

²⁴ Voir Robert Hetzron (ed.), *The Semitic Languages*, London & New York, Routledge, 1997, et notamment : ch. 7 (araméen), ch. 9 (ancien hébreu), ch. 11 (arabe classique), ch. 14 (dialectes arabes, maltais), ch. 15 (hébreu moderne), ch. 16 (araméen moderne).

²⁵ Voir Sanford B. Steever (ed.), *The Dravidian Languages*, London & New York, Routledge, 1997, et notamment le ch. 4 sur le tamoul.

apparaît bien avec les pronoms personnels de 3^{ème} personne du singulier : *ivan, ivaru, avan, avaru* (cf. anglais *he*), *iva, ivanga, ava, avanga* (cf. anglais *she*), *idu, adu* (cf. anglais *it*)²⁶.

Le chinois, langue isolante²⁷, a un système de spécificatifs et de classes qui a été mentionné plus haut. Dans ce cadre, le genre sémantique (masculin ~ féminin) peut au mieux recevoir une expression lexicale (composition). Mais là encore, l'essentiel est que le genre sémantique en chinois ne présente guère de différence avec ce qu'on rencontre dans les langues indo-européennes.

On le voit, les principes qui régissent le genre transcendent largement à la fois les regroupements géographiques ou culturels et les familles historiques.

Or, dans la mesure où tout laisse supposer que les représentations du genre diffèrent néanmoins d'une culture à l'autre, on peut en conclure que la théorie de la relation étroite entre système linguistique et représentation du monde est elle-même invalidée, au moins pour ce qui est du genre, que le classement se fasse sur la base d'une opposition géographique, culturelle, ou sur une autre base.

La femme et l'homme

Il est intéressant de passer à présent à l'analyse plus spécifique du fonctionnement des termes qui désignent la femme et l'homme. On glisse ainsi d'une étude macrostructurelle à une étude d'un cas plus précis : la désignation du genre humain et du genre de l'humain.

Commençons donc par les mots *femme* et *homme* en français. On se trouve en fait ici non pas face à deux types de signification, mais à trois : on peut spécifier que l'individu est de genre naturel féminin et l'on dit *femme* ; on peut aussi spécifier qu'il est de genre naturel masculin, et l'on dit *homme* ; mais on peut enfin estimer non pertinent ou inutile de spécifier le genre naturel, en faisant référence à une personne, ou à l'espèce dans un emploi générique, et on dit encore *homme*. Le problème est dû au fait que le système du français, peu original à cet égard, ne nous offre dans ce cas que deux possibilités sur le plan de l'expression, alors que l'on a trois types de signification. On voit que le terme grammaticalement masculin est utilisé non pas pour une, mais pour deux significations.

²⁶ Les différentes formes pour un même genre sont dues à d'autres oppositions : informel ~ formel, et proche ~ éloigné (informel : *ivan, iva, avan, ava* ; formel : *ivaru, ivanga, avaru, avanga* ; proche : *ivan, ivaru, iva, invanga* ; éloigné : *avan, avaru, ava, avangu*).

²⁷ Dans une langue isolante, les mots sont invariables.

C'est le moment de faire intervenir la théorie de la marque²⁸, ou théorie de la complexité, afin d'essayer de mettre de l'ordre dans le débat. Cette théorie postule que dans une opposition dichotomique, l'un des deux termes de l'opposition est plus complexe que l'autre.

La complexité dans le domaine du lexique ou de la grammaire concerne soit le plan de l'expression, soit le plan conceptuel, soit les deux, avec en général harmonie entre les deux plans dans ce dernier cas de figure. L'anglais illustre bien la différence de complexité observable sur le plan de l'expression, avec *man* et *woman*, le second étant plus complexe²⁹. La théorie suppose ici que la complexité morphologique va de pair avec la complexité conceptuelle. Autrement dit, on observe une relation d'iconicité.

On peut à présent se demander quelle était la situation à l'époque classique dans la *Romania*. Là où le français ne dispose plus que de deux termes (la dyade *homme* ~ *femme*), le latin classique disposait de trois : *homo* ~ *vir* ~ *mulier*. Rappelons que le premier nom de cette triade, *homo*, servait à désigner l'humain dans son ensemble, indépendamment du genre naturel ; le second, *vir*, servait à désigner un humain de genre naturel masculin, et le troisième, *mulier*, faisait référence à un humain de genre naturel féminin. Il existe donc bien des situations linguistiques dans lesquelles on dispose de trois termes pour les trois significations signalées³⁰.

Tout peut sembler parfait et « équitable » dans la situation latine. En réalité, le devenir de la langue, la fragmentation dialectale et la transformation progressive du latin parlé en langues romanes, montrent que les choses ne devaient pas être aussi simples dans la langue de César et de Cicéron. Tout d'abord, le terme servant à désigner l'être humain sans avoir à spécifier son genre naturel, *homo*, est de genre lexico-grammatical masculin, autrement dit le genre non marqué. À ce stade, l'idiome dispose non pas de trois termes de trois genres distincts, dont l'un serait un genre indifférencié, mais de deux termes de genre lexico-grammatical masculin (*homo*, *vir*), contre un seul de genre lexico-grammatical féminin (*mulier*) : comme on ne peut manquer d'identifier tant soit peu, dans le cas des êtres animés, notamment humains, le genre lexico-grammatical et le genre sémantique, le masculin lexico-grammatical se trouve ici à son avantage, qu'on le veuille ou non, avec deux termes contre un seul pour le féminin, même si l'un des deux termes masculins est censé désigner tout humain, indépendamment du sexe.

²⁸ Sur la théorie de la marque (angl. *markedness*) en phonologie, voir par exemple Roman Jakobson & Linda Waugh, *The Sound Shape of Language*, Brighton, Harvester Press, 1979, *passim*. Historiquement, cette théorie a d'abord été exploitée en phonologie, mais elle peut être appliquée à d'autres domaines linguistiques : morphologie, lexique, syntaxe, sémantique grammaticale.

²⁹ Le mot *woman* vient du vieil anglais /wi:f-man/, /wi:f/ donnant en anglais moderne *wife*, prononcé /waif/, et qui ne signifie plus que femme au sens restreint d'épouse.

³⁰ C'est aussi le cas en allemand, avec *Mensch* (« homme », sans précision de genre naturel) *Mann* (« homme », au sens masculin sur le plan naturel/conceptuel), *Frau* (« femme »).

Le développement de la langue montrera que le ver était effectivement dans le fruit. En effet, c'est le latin *homo* qui va donner *homme* en français, terme ambivalent (individu de sexe masculin ou être humain en général), tandis que *vir* ne laissera de traces que dans des dérivés savants (*viril*, *virilité*), au sémantisme très spécifique³¹. Si les choses ont évolué ainsi, c'est qu'à date ancienne, *homo* devait déjà être senti, même lorsqu'il était censé faire référence à l'humain sans spécification sexuelle, comme désignant de manière prototypique des individus de genre naturel masculin, et secondairement seulement des individus de genre naturel féminin, au mieux associés aux individus mâles. En somme, le pluriel *homines* désignant des humains (en général), faisait sans doute référence, dans l'esprit des locuteurs latinophones, soit à des individus masculins, soit à un ensemble masculin/féminin (à dominante masculine), mais en tout cas jamais uniquement à des femmes : qu'on le regrette ou non, cette asymétrie est à la fois incontestable et édifiante, car histoire linguistique et histoire culturelle se rencontrent sur ce terrain, et s'y rejoignent.

Finalement, le genre lexico-grammatical, quelles que soient les vertus de la théorie de la marque, avait donc bien favorisé le masculin sémantique au détriment du féminin.

La morale de tout cela est que l'emploi prétendument sans spécification sexuelle favorise néanmoins le genre naturel qui correspond en principe au genre lexico-grammatical non marqué, autrement dit le masculin. Il n'existe finalement aucun emploi « innocent » du genre lorsque des êtres animés, et plus particulièrement humains, sont impliqués³².

Lexicographie et idéologie

On observe que les œuvres des lexicographes eux-mêmes nous font découvrir la nature et l'évolution des représentations et des idéologies. À cet égard, il est intéressant de comparer les définitions du mot *femme* dans le Petit Larousse à un siècle d'écart. Dans l'édition de 1906, la femme est définie comme « Compagne de l'homme ; épouse » ; une centaine d'années plus tard, on lit pour la même entrée : « Être humain du sexe féminin ». Tout espoir n'est pas perdu, même si les détails sont moins encourageants. On lit en effet à l'entrée lexicale « femme » de l'édition de 2010 que « La loi salique excluait les femmes de la possession de la terre ». Cet exemple trahit-il une nostalgie pour des temps anciens plus favorables aux hommes ? On lit aussi que la quatrième acception de ce lexème « femme » est « Adulte du sexe féminin considéré par rapport à ses qualités, ses défauts, ses activités, ses

³¹ Le mot *virtu* est aussi à relier à *vir*. Quant à *femme*, il vient évidemment de *femina*, et non de *mulier*, qui a laissé des traces en italien (*moglie*) ou en castillan (*mujer*), par exemple.

³² Le genre grammatical des noms d'animaux épécènes n'est pas sans influencer la représentation que l'on se fait de l'animal désigné par le nom : voir plus loin.

origines, etc. », avec comme illustration la « Femme au foyer ». Il est intéressant de constater que si l'entrée lexicale « homme » énonce certes une acception analogue, mentionnant « qualités et défauts propres », l'adulte mâle a droit en plus, privilège qui lui est accordé, à une acception spécifique réservée à ses seules qualités : « Être humain de sexe masculin considéré du point de vue des qualités attribuées communément à son sexe (virilité, courage, etc.). » La femme n'a pas droit à de tels égards, ni à une acception analogue, sans parler de l'écart manifeste entre le nombre de lignes consacrées à l'humain de sexe masculin, nettement plus important que pour la femme.

La problématique du genre en français

Dans le cadre défini plus haut pour le français, les noms désignant des entités asexuées se retrouvent nécessairement dans la catégorie du masculin ou du féminin sur le plan lexicogrammatical, alors même que ces catégories ne sont pas ici pertinentes du point de vue sémantique. Le système optimal face à une telle contrainte serait *a priori* le suivant :

- les noms faisant référence à des entités de sexe masculin sont de genre lexico-grammatical masculin ;
- les noms faisant référence à des entités de sexe féminin sont de genre lexico-grammatical féminin ;
- les noms faisant référence à des entités asexuées sont de genre lexico-grammatical masculin ou féminin, de manière arbitraire.

Dans un tel système, l'arbitraire serait limité aux entités asexuées. Le problème est que ce système représente un idéal, mais n'est pas tout à fait conforme à la réalité. Il existe en effet plusieurs facteurs qui éloignent le système réel du français de cette situation présumée idéale.

Le premier tient à la théorie de la marque, dont les principes ont été rappelés plus haut : la conséquence de ces principes est qu'un terme de genre lexico-grammatical masculin pourra désigner une entité sexuée, envisagée comme masculine, et que le même terme pourra désigner toute entité de la catégorie sexuée concernée, mais sans prendre en compte son appartenance sexuelle. On retrouve ici une asymétrie entre masculin et féminin. Ainsi, le mot *ouvriers* pourra désigner aussi bien des ouvriers que des ouvrières.

Le second est le fait qu'au pluriel, en cas de mixité, le masculin, c'est bien connu, « l'emporte » sur le féminin : ainsi, l'expression nominale *les ouvriers* pourra être employée pour désigner un ensemble de femmes et d'hommes.

Le troisième facteur est l'existence de termes épïcènes de genre lexico-grammatical masculin, désignant aussi bien l'entité de sexe masculin que celle de sexe féminin. Ainsi, on

peut être membre d'un comité, mais que l'on soit homme ou femme, le nom *membre* reste de genre grammatical masculin.

Dans tous les cas qui viennent d'être mentionnés, le masculin sort en quelque sorte « vainqueur ». Mais il existe aussi une autre situation, inverse de la précédente, dans laquelle le nom est de genre lexico-grammatical féminin, et peut désigner un individu de l'un ou l'autre sexe : c'est le cas de *sentinelle*, par exemple, terme qui sera d'ailleurs employé pour une référence à un homme plus facilement qu'à une femme.

Pour ce qui est des termes de genre lexico-grammatical masculin pouvant désigner aussi bien des femmes que des hommes, une évolution récente, qui est l'une des rares conséquences de l'interventionnisme linguistique, permet que la plupart des noms désignant des fonctions, tels que *ministre*, s'emploient désormais aussi bien au féminin qu'au masculin : *le ministre, la ministre*, alors qu'à une époque encore peu éloignée la norme exigeait *Madame le ministre*, par exemple, et l'expression *la ministre* était volontiers stigmatisée.

Malgré cette évolution, il existe un autre facteur qui assure une domination du genre lexico-grammatical masculin dans l'usage. C'est le fait que lorsque l'on pose une notion avant d'avoir à déterminer le genre des entités auxquelles elle s'applique, le masculin semble bien s'imposer d'office. Soit un homme et une femme se concertant pour savoir qui commence à parler ou à jouer, etc. L'un ou l'autre dira : « Lequel de nous deux commence ? », avec le pronom interrogatif masculin. Dans cet esprit, force est de constater l'asymétrie du chiasme apparent que forment les deux énoncés suivants, de type équatif :

(a) « Le ministre de la santé est une femme. »

(b) « La ministre de la santé est un homme ».

Chacun conviendra que les effets de sens ne sont pas croisés, contrairement à la forme de l'expression. L'énoncé (a) s'impose si la personne exerçant la fonction ministérielle est une femme, alors que l'énoncé (b) est incongru. Ici, la théorie de la marque impose à nouveau sa tyrannie : en (a), *le ministre*, sujet de genre masculin, donc non marqué, peut désigner aussi bien une femme qu'un homme, laissant la porte ouverte sans contradiction à une spécification féminine dans le prédicat ; en revanche, en (b), *la ministre*, forme innovante, mais néanmoins soumise aux principes, est marquée : en tant que telle, cette forme désigne nécessairement une femme, ce qui entraîne inéluctablement un effet de sens de type rhétorique à l'audition du prédicat. En effet, celui-ci énonce une spécification masculine, donc contraire au genre du sujet. Le récepteur du message, qui croit en principe à la rationalité à l'énonciateur, se dirigera, en empruntant la voie de la rhétorique, vers une interprétation réconciliant les

contraires en (b), alors que l'énoncé (a) était exempt d'une telle contrariété dans l'interprétation préférentielle, celle qui vient en premier à l'esprit.

Ces exemples montrent bien que les innovations, même si elles représentent incontestablement un progrès considérable, ne résolvent pas tout : on peut bien réformer quelques bribes de lexique, mais on ne réforme pas la langue en profondeur par décret.

Par ailleurs, il semble bien que les locuteurs francophones aient une tendance très forte à un certain type de fonctionnement au niveau psycholinguistique. L'association prototypique correspond à un principe de recherche de congruence entre genre lexico-grammatical et genre sémantique (naturel) dans le cas des entités sexuées. Ainsi, le puma sera vu plus facilement comme un mâle, et la girafe comme une femelle. Ce principe de congruence donne même des résultats étonnants dans le cas de termes synonymes désignant des animaux, l'un masculin et l'autre féminin, tels que *léopard* et *panthère*. Il s'agit dans les deux cas du même animal, les deux noms étant interchangeable, sauf dans le cas d'un individu au pelage noir, auquel on applique uniquement le nom *panthère*. Peu importe que notre locuteur sache ou non qu'il s'agit de deux désignations distinctes du même animal : il associera plus volontiers *léopard* au sexe masculin et *panthère* au sexe féminin³³.

Dans la même logique, alors même que chacun sait que les girafes sont de sexe masculin ou féminin, il semble bien qu'une majorité de locuteurs pensent plus spontanément à une femelle lors de l'emploi du nom *girafe*, en raison du genre lexico-grammatical féminin de ce lexème. Si c'est exact, cela ne remet certes pas en question l'idée que le genre en français est lexico-grammatical et largement immotivé, mais fait douter de la validité d'une analyse qui poserait que le genre lexico-grammatical et le genre sémantique sont des catégorisations largement indépendantes. Bien au contraire, la complexité du système français est en effet en partie liée à cette interdépendance relative de deux catégorisations appartenant à des plans différents : classes lexico-grammaticales d'un côté, classes sémantiques de l'autre. L'indépendance des deux plans n'est à peu près nette que dans le cas des inanimés les plus typiques, comme *chaise*, *crayon*, *caillou*, etc. Mais même cela n'empêche pas que certaines réalités naturelles, non sexuées, seront facilement l'objet d'un rapprochement entre genre lexico-grammatical et genre sémantique naturel, poétique ou métaphorique : comme le dit la chanson, le soleil a rendez-vous avec la lune, alors que d'autres langues offrent le schéma inverse : l'allemand a *die Sonne* (féminin), pour « le soleil », et *der Mond* (masculin), pour « la lune », et le dyirbal, langue aborigène d'Australie en voie d'extinction, fait précéder

³³ D'ailleurs, lorsque le nom *panthère* est métaphoriquement appliqué à un être humain, c'est forcément à une femme.

chacun des deux noms d'un spécificateur de genre : le soleil se dit *balan garri*, *balan* signifiant « féminin », et la lune se dit *bayi gagara*, *bayi* signifiant « masculin »³⁴.

Conclusion

Dans le domaine du genre, l'on aurait tort de séparer radicalement le plan culturel ou idéologique du plan linguistique. En effet, d'un côté, les résultats de l'analyse macro-structurale laissent supposer l'existence de principes généraux à l'œuvre dans une sorte de grammaire universelle du genre, ainsi qu'une relative indépendance du genre lexicogrammatical et du genre sémantique dans des langues comme le français, par exemple ; mais d'un autre côté, l'étude diachronique, l'analyse micro-structurale, ainsi que l'observation des habitudes lexicographiques, révèlent des relations dialectiques entre les deux plans, ainsi qu'entre langues et mentalités.

Cela signifie que même si les principes universels sous-jacents régissant le fonctionnement des langues vont à l'encontre à la fois du relativisme et de l'idée d'une interdépendance forte entre structures grammaticales et représentations culturelles, le genre est l'une des rares catégories qui se situent manifestement à l'intersection de ces axes. Toutefois, empressons-nous d'ajouter qu'il semble bien que l'interaction se caractérise ici par l'unité plutôt que par la diversité.

Il s'agit en tout cas d'histoires complexes, à la fois différentes et reliées, mais au sein de cette complexité, une chose est certaine pour ce qui nous concerne aujourd'hui : les archaïsmes se transmettent aussi facilement que les innovations.

Références

- Darrobers, Roger & Planes, Xiaohong Xiao, *Éléments fondamentaux de la phrase chinoise*, Paris, éd. You-Feng, 1998.
- Dixon, Robert M. W., *The Languages of Australia*, Cambridge University Press, 1980.
- Doke, Clement Martin et al., *English-Zulu, Zulu-English Dictionary*, Johannesburg, Witwatersrand University Press, 1990.
- Haudry, Jean, *L'indo-européen*, Paris, P.U.F., 1979.
- Hetzron, Robert (ed.), *The Semitic Languages*, London and New York, Routledge, 1997.
- Humbert, Jean, *Syntaxe grecque*, Paris, Klincksieck, 1945, 1960, 1972, 2004.
- Jakobson, Roman & Waugh, Linda, *The Sound Shape of Language*, Brighton, Harvester Press, 1979.
- Martinet, André, *Évolution des langues et reconstruction*, Paris, P.U.F., 1975.

³⁴ Voir Robert M. W. Dixon, *The Languages of Australia*, Cambridge University Press, 1980, p. 273-274.

- Martinet, André, *Des steppes aux océans*, Paris, Payot & Rivages, 1986, 1994.
- Matthews, Stephen & Yip, Virginia, *Cantonese Grammar*, London and New York, Routledge, 1994.
- Norman, Jerry, *Chinese*, Cambridge University Press, 1988.
- Roche, Philippe, *Grammaire active du chinois*, Paris, Larousse, 2007.
- Saussure, Ferdinand de, *Cours de linguistique générale*, Paris, Payot, 1980 (initialement publié en 1916).
- Steever, Sanford B. (ed.), *The Dravidian Languages*, London and New York, Routledge, 1997.
- Wald, Benji, "Swahili and the Bantu Languages", in Bernard Comrie (ed.), *The World's Major Languages*, London & Sydney, Croom Helm, 1987, p. 991-1014.