

Mémo RLF
Traitement lexicographique
des correspondances $N_{\text{masc}} \sim N_{\text{fém}}$ en français :
la réponse du berger à la bergère

Candice Delaite⁽¹⁾, Alain Polguère⁽²⁾

⁽¹⁾ATILF CNRS, ⁽²⁾Université de Lorraine & ATILF CNRS

26 avril 2013

Table des matières

1	Problème traité	2
1.1	Paires $N_{\text{masc}} \sim N_{\text{fém}}$ fondées sur la différence de sexe	2
1.2	Dérivation plutôt que flexion	3
1.3	Hypothèse : deux types de paires $N_{\text{masc}} \sim N_{\text{fém}}$	4
2	Quasi-synonymie relative au sexe	4
2.1	Caractérisation	4
2.2	Générique masculin ou féminin	6
3	Opposition contrastive relative au sexe	6
4	Implications sur la structure du RLF	8
	Remerciements	9
	Bibliographie	9

Résumé

Nous examinons les paires de lexies nominales du français du type *berger* \sim *bergère*, *chien* \sim *chienne*, *étalon* \sim *jument*, *fiancé* \sim *fiancée*. . . Ces paires sont examinées sous l'angle de leur modélisation lexicographique dans le cadre du projet de construction du Réseau Lexical du Français (RLF). Nous détaillons tout d'abord le problème traité (section 1), en identifiant notamment deux types principaux de relations sémantiques mises en jeu dans de telles paires : la quasi-synonymie relative au sexe (traitée en détail dans la section 2) et l'opposition contrastive relative au sexe (traitée dans la section 3). En conclusion (section 4), nous examinons les implications de notre modélisation sur la structure du RLF.

1 Problème traité

1.1 Paires $N_{\text{masc}} \sim N_{\text{fém}}$ fondées sur la différence de sexe

Nous nous intéressons dans ce court texte à la description lexicographique des paires nominales du français que nous appelons $N_{\text{masc}} \sim N_{\text{fém}}$ et qui possèdent les quatre caractéristiques suivantes :

1. les lexies N_{masc} et $N_{\text{fém}}$ dénotent toutes deux un être animé, du même type, modulo une différence de sexe ;
2. N_{masc} est du genre masculin et dénote un être animé qui est, nécessairement ou par défaut, de sexe masculin ;
3. $N_{\text{fém}}$ est du genre féminin et dénote un être animé qui est, nécessairement ou par défaut, de sexe féminin ;
4. N_{masc} et $N_{\text{fém}}$ sont généralement, mais non nécessairement, en relation morphologique.

Voici quelques exemples : *berger* \sim *bergère*, *hôte* \sim *hôtesse*, *nageur* \sim *nageuse*, *chien* \sim *chienne*, *étalon* \sim *jument*, *fiancé* \sim *fiancée*...

Nous examinons ces paires lexicales dans le cadre du projet lexicographique de construction du Réseau Lexical du Français ou RLF (Lux-Pogodalla et Polguère, 2011; Gader et al., 2012). Le RLF est un modèle du lexique français élaboré selon les principes de la Lexicologie Explicative et Combinatoire (Mel'čuk et al., 1995) et qui a la particularité d'être formellement un graphe lexical multidimensionnel, plutôt qu'un « texte dictionnaire ». L'essentiel de la structuration du RLF repose sur le tissage de liens de fonctions lexicales (Mel'čuk, 1996).

Puisque notre perspective est lexicographique, il est utile de commencer par résumer brièvement la façon dont les dictionnaires « classiques » traitent les paires $N_{\text{masc}} \sim N_{\text{fém}}$, en nous focalisant sur les paires de lexies morphologiquement liées.

Les trois principaux dictionnaires de référence du français – *Petit Robert*¹, *Petit Larousse*² et *Trésor de la Langue Française*³ – traitent les lexies $N_{\text{masc}} \sim N_{\text{fém}}$ morphologiquement liées de façon hétérogène. En effet, dans certains cas, il existe une entrée pour chaque lexie de la paire – par exemple, MARQUIS \sim MARQUISE ou POÈTE \sim POÉTESSE – et dans d'autres, les deux lexies sont regroupées sous une même entrée – par exemple, COIFFEUR \sim COIFFEUSE, décrites sous une entrée commune nommée « **coiffeur, -euse** ». Ces dictionnaires semblent converger dans leurs choix descriptifs et on peut généralement trouver une explication logique à ceux-ci. Ainsi, dans le cas des exemples cités :

- l'entrée séparée de MARQUISE est quasiment imposée aux dictionnaires par la polysémie qui s'est développée autour du nom féminin – *marquise* (*d'une gare*), *marquise* (*au chocolat*) –, sans correspondance dans le vocable masculin ;
- le suffixe dérivationnel *-esse*⁴ a un statut particulier en morphologie française qui fait que POÉTESSE apparaît comme morphologiquement plus « éloigné » de POÈTE que COIFFEUSE ne l'est de COIFFEUR.

Le traitement lexicographique des dérivations sémantiques relatives au sexe dépend bien évidemment des caractéristiques grammaticales de chaque langue. Ainsi, comme il n'existe pas de flexion en genre en anglais, le problème de la nature morphologique de la relation entre la source et la cible de telles dérivations ne se pose pas. De plus,

1. Édition cédérom de 2011.

2. Consultation en ligne : <http://www.larousse.com/fr/dictionnaires/francais-monolingue>.

3. Consultation en ligne : <http://atilf.atilf.fr>.

4. Il s'agit d'un suffixe archaïque plutôt rare dans le lexique du français contemporain ; il ne s'emploie quasiment plus pour créer des néologismes. Voir à ce propos ce que dit le TLF dans sa notice étymologique, sous l'entrée *-ESSE*².

de telles paires – fondées sur l’usage du suffixe angl. *-ess* (emprunté au français) – sont relativement peu nombreuses. Pour ces raisons, les dictionnaires de l’anglais créent toujours (nous semble-t-il) deux entrées séparées, pour le nom dénotant un être animé de sexe masculin et pour le nom dénotant l’être animé correspondant de sexe féminin – voir, par exemple, LION ‘lion’ ~ LIONESS ‘lionne’ dans *American Heritage*⁵, *Longman Dictionary of Contemporary English*⁶ ou *Cobuild*⁷.

Le cas de l’espagnol est plus intéressant ici, puisque que les noms ont un genre grammatical dans cette langue, dont la grammaire comporte une flexion en genre pour les adjectifs, les participes et les articles. On y trouve par ricochet, comme en français, beaucoup de paires $N_{\text{masc}} \sim N_{\text{fém}}$. Tout comme les dictionnaires du français, le *Diccionario de la lengua español* (Real Academia Española)⁸ et le *Diccionario de uso del español* Moliner (2000) proposent un traitement hétérogène des $N_{\text{masc}} \sim N_{\text{fém}}$: deux entrées dissociées pour MODISTO ‘couturier’ ~ MODISTA ‘couturière’, mais une seule entrée pour PASTOR ‘berger’ et PASTORA ‘bergère’ intitulée « **pastor, ra** » dans le premier dictionnaire et « **pastor, -a** » dans le second. Le cas de l’espagnol est cependant très différent du français dans la mesure où chaque nom commun de cette langue, dans les cas standard, porte un suffixe qui marque le genre grammatical – cf. Mel’čuk (à paraître) pour une analyse de la nature morphologique des suffixes nominaux de genre de l’espagnol.

Dans le cadre du RLF, nous adoptons un traitement tout à fait systématique des paires $N_{\text{masc}} \sim N_{\text{fém}}$ morphologiquement liées, qui repose notamment sur la nature dérivationnelle du lien morphologique qu’elles mettent en jeu. Examinons brièvement ce point.

1.2 Dérivation plutôt que flexion

Il est clair que, du point de vue de la structure des signifiants, une paire comme *berger* ~ *bergère* présente toutes les apparences de la flexion, la forme *bergère* apparaissant comme un « féminin de *berger* ». Cependant, comme l’explique Mel’čuk (2000) – dont nous reprenons dans cette section l’essentiel de l’argumentation –, il est impossible de considérer que le mécanisme morphologique mis en œuvre ici en est un de flexion, pour au moins cinq raisons.

1. Il faudrait postuler pour le français une catégorie flexionnelle de genre nominal exclusivement applicable aux noms dénotant des êtres animés sexués, alors que la flexion type s’applique de façon globale à toutes les lexies d’une partie du discours donnée (comme le genre des adjectifs français, par exemple).
2. La flexion hypothétique aurait des manifestations morphologiques extrêmement irrégulières – PROFESSEUR ~ PROFESSEURE, mais COIFFEUR ~ COIFFEUSE – et non systématiques puisqu’elle ne s’appliquerait tout simplement pas dans tous les cas – cf. POISSON_(N, masc) qui n’a pas de contrepartie *POISSON_(N, fém) ou *POISSONNE.
3. Pour ces raisons, il n’est pas rare que le locuteur natif hésite sur le choix des termes à utiliser, ce qui est étonnant pour une « flexion »; ajoutons que les variations diatopiques que l’on constate pour les paires $N_{\text{masc}} \sim N_{\text{fém}}$ sont perçues comme relevant de la variation lexicale plutôt que grammaticale – cf. ÉCRIVAINNE, presque systématiquement employé au Québec, *vs* ÉCRIVAIN_(N, fém), beaucoup plus utilisé en France.
4. Comme nous l’avons mentionné dans la section 1.1, chacun des éléments d’une paire $N_{\text{masc}} \sim N_{\text{fém}}$ tend à développer sa propre polysémie; on constate donc la présence de vocables masculins *vs* féminins structurés de façon autonome, ce qui contredit une modélisation flexionnelle du phénomène en question.

5. Consultation en ligne : <http://www.ahdictionary.com>.

6. Consultation en ligne : <http://www.ldoceonline.com>.

7. Consultation en ligne : <http://dictionary.reverso.net/english-cobuild/teach>.

8. 20^e édition, consultation en ligne : <http://www.rae.es/rae.html>.

5. Finalement, et c'est fondamental, la description dérivationnelle plutôt que flexionnelle des paires $N_{\text{masc}} \sim N_{\text{fém}}$ du français se fait de façon simple et directe. La seule justification que l'on puisse trouver à l'approche courante des dictionnaires, qui tendent à modéliser ces paires comme s'il s'agissait de flexions (section 1.1 ci-dessus), est l'économie de pages dans les volumes imprimés.

Ces observations amènent fort logiquement à considérer que nous sommes en présence, avec les paires $N_{\text{masc}} \sim N_{\text{fém}}$ morphologiquement liées, de cas de **dérivations** où une lexie nominale masculin dénotant un être animé de sexe masculin est utilisée comme source pour dériver une lexie nominale féminin dénotant l'être animé de sexe féminin correspondant.

1.3 Hypothèse : deux types de paires $N_{\text{masc}} \sim N_{\text{fém}}$

Maintenant qu'il est établi que les paires nominales de type *boulangier* \sim *boulangère*, *fil* \sim *fille*, *chat* \sim *chatte*, etc., correspondent bien à des cas de dérivations morphologiques et ne relèvent pas d'un mécanisme de flexion, nous allons identifier les types majeurs de telles paires dérivationnelles et indiquer comment chaque type doit être modélisé dans le RLF.

Rappelons que nous nous intéressons en fait plus généralement ici à des liens de dérivations sémantiques non nécessairement morphologiques et que nous considérons ensemble aussi bien les paires morphologiquement liées mentionnées ci-dessus que des paires comme FRÈRE \sim SŒUR ou ÉTALON \sim JUMENT.

Nous nous situons en effet dans la perspective lexicographique de la construction du RLF et du tissage de sa structure au moyen de liens de fonctions lexicales. Dans un tel contexte, il est essentiel que les paires lexicales $N_{\text{masc}} \sim N_{\text{fém}}$ soient décrites de façon homogène, qu'elles mettent ou non en jeu une relation morphologique.

Il existe deux types majeurs de relations entre une lexie nominale masculin dénotant un être animé de sexe masculin et une lexie nominale féminin dénotant l'être animé de sexe féminin correspondant. Ces relations peuvent être exemplifiées à l'aide de deux paires :

1. la paire AVOCAT \sim AVOCATE, qui relève, selon nous, d'une relation de nature synonymique ;
2. la paire ÉTALON \sim JUMENT, qui relève plutôt d'une relation de nature contrastive.

Nous allons maintenant examiner successivement ces deux cas de figure.

2 Quasi-synonymie relative au sexe

2.1 Caractérisation

Dans la paire AVOCAT \sim AVOCATE, comme dans tous les cas de noms de professions ou d'activités, une des deux lexies – le nom masculin ou le nom féminin (voir section 2.2, ci-dessous) – sert de générique.

Ainsi :

- La forme *avocat* en (1a) ci-dessous dénote nécessairement un individu mâle alors que le sexe n'est pas déterminé en (1b).

- (1)
 - a. *Le jeune avocat s'est avancé résolument vers les journalistes.*
 - b. *Il a demandé la présence d'un avocat.*

- Le pluriel *avocats* en (2a) peut désigner un groupe incluant hommes et femmes et (2b) peut tout à fait être énoncé par une femme.

- (2) a. *Tous les avocats du barreau parisien ont protesté.*
 b. *Je veux devenir avocat.*

– Une femme peut d’ailleurs en fait tout aussi bien dire (2b) que (3).

- (3) *Je veux devenir avocate.*

Il y a donc un débalancement sémantique entre les deux lexies AVOCAT et AVOCATE :

- (4) a. ‘avocat’ = ‘individu (de sexe masculin) qui exerce une profession consistant à [...] ou profession correspondante’
 b. ‘avocate’ = ‘individu de sexe féminin qui exerce une profession consistant à [...] ou profession correspondante’

En d’autres termes, ces lexies sont toutes deux sémantiquement ambivalentes⁹, entre dénotation d’un individu ou d’une profession, mais pas de la même façon. Que ce soit pour dénoter un individu ou la profession qu’il exerce, AVOCATE implique le sexe féminin, ce qui n’est pas le cas de AVOCAT pour ce qui est du sexe masculin. La composante ‘de sexe masculin’ de la définition de ce second lexème doit être considérée comme faible (Mel’čuk et al., 1995, p. 95) ; c’est ce caractère sémantiquement faible que signalent les parenthèses dans la glose (4a) ci-dessus.

Ce rapport sémantique complexe fait que les lexies formant des paires de ce type ne sont pas perçues comme étant en opposition vis-à-vis de la dénotation du sexe. Les éléments de la paire AVOCAT ~ AVOCATE ne s’opposent pas en tant que désignant un « mâle » *vs* une « femelle » correspondante. À cause de cela, nous considérons que de telles lexies sont des quasi-synonymes. Nous appelons ce type de relation la *quasi-synonymie relative au sexe*.

Si nous prenons en compte le débalancement sémantique identifié plus haut, il nous faut distinguer les deux cas ci-dessous.

1. Un locuteur qui dit (5a) au lieu de l’énoncé typique (5b) choisit d’être plus spécifique quant à la désignation du sexe.

- (5) a. – *Je réclame une avocate !*
 b. – *Je réclame un avocat !* [homme ou femme]

Nous encodons donc le lien de dérivation sémantique AVOCAT→AVOCATE par la fonction lexicale **Syn**^{sex}, pour ‘synonyme plus riche vis-à-vis du sexe (lat. *sexus*)’.

2. Un locuteur qui dit (5b) plutôt que (5a), choisit d’être moins spécifique quant à la désignation du sexe. Nous encodons donc le lien de dérivation sémantique AVOCATE→AVOCAT par la fonction lexicale **Syn**_C^{sex}, pour ‘synonyme potentiellement moins riche vis-à-vis du sexe’. Nous parlons ici de synonyme **potentiellement** moins riche, car le locuteur peut bien entendu aussi utiliser *avocat* pour désigner spécifiquement un individu de sexe masculin.

Notons que les deux fonctions lexicales qui viennent d’être présentées correspondent à deux relations réciproques :

$$\mathbf{Syn}_{\supset}^{\text{sex}}(L_1) = L_2 \text{ entraîne que } \mathbf{Syn}_{\subset}^{\text{sex}}(L_2) = L_1, \text{ et réciproquement.}$$

Dans le domaine des professions, activités, etc., on trouve bien entendu un nombre considérable de paires fondées sur ces synonymies approximatives. Citons ainsi les types de dénnotations suivants :

9. Sur l’ambivalence sémantique et son traitement lexicographique, voir Milićević et Polguère (2010). Notons que l’ambivalence sémantique étant ici systématique en français – tout nom dénotant un individu exerçant une certaine profession peut aussi dénoter la profession elle-même –, Milićević et Polguère (2010) proposent de ne pas l’expliciter dans la structure de la définition et de simplement formuler le genre prochain de ce type de lexie selon le patron ‘individu qui exerce une certaine profession’.

- personnes ayant un titre ou fonction officiel comme ROI \sim REINE, MARQUIS \sim MARQUISE, PRÉSIDENT \sim PRÉSIDENTE ou PAPE \sim PAPESSE ;
- personnes ayant une certaine nationalité, religion, etc., comme FRANÇAIS_(N) \sim FRANÇAISE, JUIF_(N) \sim JUIVE ou BOUDDHISTE_(N, masc) \sim BOUDDHISTE_(N, fém) ;
- personnes éprouvant un sentiment ou ayant une certaine caractéristique de comportement comme AMOUREUX_(N) \sim AMOUREUSE ou LÂCHE_(N, masc) \sim LÂCHE_(N, fém) ;
- animaux, comme CHIEN \sim CHIENNE ou ÉLÉPHANT \sim ÉLÉPHANTE.

Attention à ce dernier cas, cependant ! Nous ne prétendons aucunement que toutes les paires de noms relevant de ces désignations sont en relation de **Syn**^{sex} *vs* **Syn**_c^{sex}. Dans de nombreux cas, par exemple la paire de noms d'animaux ÉTALON \sim JUMENT, nous sommes en présence d'un autre type de relation, de nature contrastive, qui sera examiné dans la section 3.

2.2 Générique masculin ou féminin

Il est important de souligner que le ratio sémantique entre les deux lexies liées par la quasi-synonymie relative au sexe, de même que la présence d'un quasi-synonyme de ce type pour un nom masculin *vs* féminin, ne sont aucunement systématiques en français, comme le montrent les deux cas de figure ci-dessous.

1. Ce peut être le nom féminin (plutôt que le nom masculin) qui, tout en dénotant un être animé de sexe féminin, sert de générique. Ainsi, OIE dénote l'animal femelle ou l'espèce, son **Syn**_c^{sex} étant JARS, qui ne peut aucunement dénoter l'espèce.
2. Un nom féminin d'animal peut être complètement neutre quant à la dénotation d'un sexe et donc n'avoir aucun pendant masculin. Tel est le cas de GIRAFE. Pour dénoter une girafe d'un sexe ou de l'autre, on dira *girafe mâle* \sim *girafe femelle*.

Certaines généralisations peuvent cependant être faites. Il nous semble ainsi que, en français, le générique est toujours le nom masculin **dans le cas de la désignation d'êtres humains**. Nous n'avons pas trouvé de contre-exemple. Même une paire comme INFIRMIER \sim INFIRMIÈRE, pour laquelle on pourrait penser que le féminin est le générique, n'échappe pas à cette règle. En effet, le féminin est plutôt la valeur par défaut pour des raisons extra-linguistiques, tenant au fait que la profession a longtemps été exercée exclusivement par des femmes. Mais *infirmière* dénote bien exclusivement un individu de sexe féminin et ne peut être utilisé comme générique ; cf. la phrase (6) ci-dessous, qui ne peut pas servir dans le cas où il est question d'un groupe de femmes et d'hommes.

(6) *Un groupe d'infirmières se tenait à l'entrée de l'hôpital.*

Contrairement à ce que l'on pourrait penser a priori, c'est bien le nom masculin qui sert aujourd'hui de générique. Ainsi, dans (7) ci-dessous, tiré du Web, il peut être question aussi bien de vêtements pour hommes que pour femmes :

(7) *L'association a ainsi acheminé du matériel médical : 140 lits médicalisés manuels ou électriques, un bloc opératoire, 4 750 paires de lunettes de vue et des vêtements d'infirmiers.*

Il existe vraisemblablement d'autres généralisations de ce type, qui pourraient être faites à propos de la quasi-synonymie fondée sur le sexe, mais nous ne disposons pas à ce jour de suffisamment de données pour les identifier.

Passons maintenant au second type de dérivation sémantique relative au sexe.

3 Opposition contrastive relative au sexe

Tel que mentionné plus haut, certaines paires $N_{\text{masc}} \sim N_{\text{fém}}$ ne relèvent pas de la quasi-synonymie. Il serait très étrange de dire, par exemple, que ÉTALON et JUMENT

sont des quasi-synonymes, alors qu'ils sont plutôt perçus comme sémantiquement en opposition : l'étalon est le cheval mâle, par opposition à la jument, le cheval femelle.

Quelle fonction lexicale utiliser pour rendre compte du lien sémantique mis en jeu dans les paires lexicales de ce type, puisque le recours à **Syn** est exclu ? Il existe bien entendu la fonction lexicale « de contrastivité » **Contr** ; celle-ci n'est cependant pas en jeu ici, pour des raisons que nous allons expliquer. Rappelons pour ce faire que **Contr** se caractérise de la façon suivante :

Contr(L_1) = L_2 si l'on trouve dans la langue des expressions phraséologiques mettant en opposition contrastive les deux lexies L_1 et L_2 . Par exemple, **Contr**(*noir*) = *blanc* car il existe des expressions comme (*photo en noir et blanc*, *Tout n'est pas tout blanc ou tout noir*, etc.¹⁰

Cette caractérisation montre bien que des paires comme ÉTALON ~ JUMENT, tout en présentant un contraste sémantique évident, ne relèvent pas de **Contr**. En effet, elles sont reliées directement par une opposition sémantique issue de leur définition lexicographique, et non par l'existence en langue d'expressions phraséologiques les contrastant.

Si la relation considérée ici est associée, malgré ses différences, à **Contr**, c'est parce que comme cette dernière elle se distingue radicalement de l'antonymie – fonction lexicale **Anti**. Elle ne repose pas sur une opposition sémantique relevant de la négation d'une composante de sens : comparez sémantiquement PRÉSENT_(Adj) 'qui est là' vs son antonyme ABSENT_(Adj) 'qui n'est pas là' avec ÉTALON vs JUMENT. La nature non antonymique de la relation entre ÉTALON et JUMENT apparaît clairement si l'on ébauche les définitions des deux sémantèmes qui les opposent – '[X] femelle', inclus dans le sens de JUMENT et '[X] mâle', dans le sens de ÉTALON –, en nous inspirant très librement du traitement sémantique des lexies dénotant des êtres sexués proposé par (Wierzbicka, 1972, p. 34–56) :

- '[X] femelle' = '[X] dont le sexe peut lui permettre de donner naissance'
- '[X] mâle' = '[X] dont le sexe peut lui permettre de participer au fait qu'un individu femelle de la même espèce donne naissance'

La relation qui nous intéresse étant proche – mais distincte ! – de **Contr**, elle doit en théorie être modélisée au moyen de deux fonctions lexicales non standard :

1. **Être animé femelle correspondant**(*étalon*) = *jument* ;
2. **Être animé mâle correspondant**(*jument*) = *étalon*.

Cependant, ces deux fonctions lexicales ne sont pas loin de répondre aux critères de détermination de FL standard (Polguère, 2007), notamment : elles correspondent à des liens très réguliers en français (et vraisemblablement dans toutes les langues), elles s'appliquent à un ensemble très grand et varié d'arguments, elles retournent un grand nombre de valeurs distinctes et, finalement, elles tendent à s'exprimer morphologiquement en langue.

C'est pourquoi nous avons décidé de standardiser les deux relations symétriques en question en introduisant deux fonctions lexicales normalisées dans le cadre du projet RLF :

1. **Fem** pour lat. 'feminus' ;
2. **Masc** pour lat. 'masculus'.

Les encodages **Fem** et **Masc** ont été utilisés auparavant dans le projet *DiCo* (Polguère, 2000; Mel'čuk et Polguère, 2006), dans le cadre des propositions de normalisation de certaines fonctions lexicales non standard faites par Anne-Laure Jousse (Jousse, 2010, p. 139–140). Il ne s'agissait cependant pas techniquement des mêmes **Fem** et **Masc**

10. À cause de cela, on peut dire que **Contr** n'est pas une fonction lexicale paradigmatique ordinaire, car elle ne rend pas compte d'une constante sémantique. Une lexie qui est valeur de **Contr** d'une autre lexie n'en est pas, strictement parlant, un dérivé sémantique.

dont il est fait usage dans dans le RLF puisque le DiCo les utilisait sans discrimination pour les deux familles de dérivation – synonymiques *vs* contrastives – que nous avons identifiées.

Remarquons que les paires connectées par les relations **Fem** *vs* **Masc** peuvent inter-agir dans le lexique avec une troisième lexie qui est, elle, en relation de quasi-synonymie relative au sexe avec les deux premières. Tel est le cas de CHEVAL et de sa relation avec la paire ÉTALON \sim JUMENT. La définition de CHEVAL se rapproche de celle de lexies du type AVOCAT examinées dans la section précédente. Le nom CHEVAL dénote un animal d’un certain type qui est, par défaut et non nécessairement, de sexe mâle. On obtient ainsi un système complexe de relations lexicales unissant les lexies CHEVAL, ÉTALON et JUMENT présenté dans la Figure 1 ci-dessous.

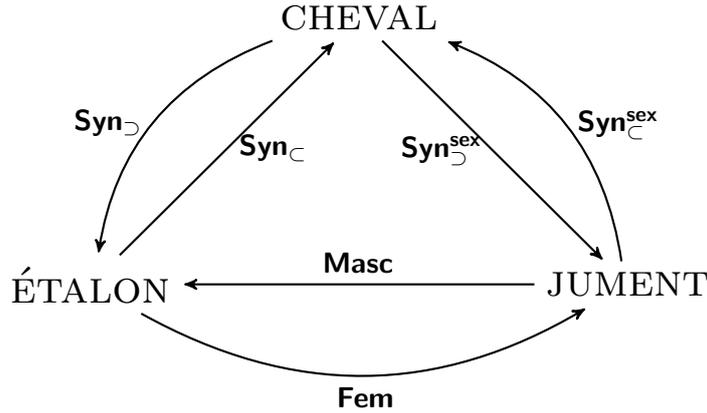


FIGURE 1 – Système de relations lexicales unissant CHEVAL, ÉTALON et JUMENT

On notera dans la figure ci-dessus l’usage des fonctions lexicales **Syn_D** et **Syn_C**, plutôt que **Syn_D^{sex}** et **Syn_C^{sex}**, pour connecter CHEVAL et ÉTALON. En effet, ces deux lexies n’entretiennent pas de différence sémantique relative au sexe ; ÉTALON est plus spécifique que CHEVAL du fait que sa définition comprend une composante sémantique ‘mâle’ qui n’est pas faible ainsi qu’une composante relative à la non-castration et à une fonction de reproduction.

Pour conclure, nous énumérons maintenant, sans développer, deux caractéristiques des oppositions contrastives relatives au sexe que nous avons pu identifier **pour le français**.

1. Les relations **Fem** et **Masc** sont bien moins présentes dans le lexique que **Syn_D^{sex}** et **Syn_C^{sex}**. Ceci s’explique assez facilement par le fait que la synonymie approximative est justement la relation la plus commune dans tout réseau lexical.
2. **Syn_D^{sex}** *vs* **Syn_C^{sex}** est souvent morphologisé, alors que **Fem** *vs* **Masc** l’est difficilement. Dans les faits, nous n’avons pas trouvé de cas clair de **Fem** morphologiquement construit sur son **Masc** (ou vice versa).

Insistons sur le fait qu’il ne s’agit ici que d’observations préliminaires et que les points mentionnés ci-dessus devraient être examinés de façon systématique.

4 Implications sur la structure du RLF

L’analyse que nous venons de présenter a deux répercussions directes sur la façon dont se développe la structure du RLF.

Premièrement, le RLF possède systématiquement des entrées lexicographiques distinctes pour les paires $N_{\text{masc}} \sim N_{\text{fém}}$. Du fait que les vocables de ces paires peuvent développer leur polysémie de façon indépendante, il ne faut jamais les fusionner sous

une même entrée, même lorsqu'ils font appel aux mêmes signifiants – cf. $L\hat{A}CHE_{(N, masc)} \sim L\hat{A}CHE_{(N, f\acute{e}m)}$. L'économie de volume d'impression n'est pas un paramètre dans la construction du RLF et nous pouvons nous permettre de coller au plus près à la réalité de la langue. De plus, seule notre approche permet d'offrir une description complète et autonome pour toute lexie – caractéristiques grammaticales, définition, dérivés sémantiques, collocatifs et exemples lexicographiques –, ce qui est bien entendu souhaitable pour l'utilisateur du RLF¹¹.

Deuxièmement, les liens paradigmatiques connectant les lexies impliquées dans le système des oppositions $N_{masc} \sim N_{f\acute{e}m}$ sont systématiquement encodés par les lexico-graphes du RLF au moyen des fonctions lexicales appropriées – **Syn**_▷^{sex}, **Syn**_◁^{sex}, **Masc** et **Fem**. Au moment où nous écrivons, l'encodage des relations lexicales du RLF relevant des paires $N_{masc} \sim N_{f\acute{e}m}$ est constitué de 1 680 liens de fonctions lexicales : 1 560 liens pour les relations **Syn**_▷^{sex} \sim **Syn**_◁^{sex} et 120 liens seulement pour les relation **Fem** \sim **Masc**. Ces statistiques confirment ce que nous avons dit plus haut, à savoir que la synonymie relative au sexe est largement plus présente dans le lexique que l'opposition contrastive correspondante.

Pour terminer, mentionnons qu'il existe des degrés dans la nature synonymique *vs* contrastive des liens unissant les lexies $N_{masc} \sim N_{f\acute{e}m}$. Ainsi, la paire FILS \sim FILLE relève de l'opposition contrastive, puisqu'aucune des deux lexies ne peut être utilisée de façon générique pour dénoter un enfant de quelqu'un sans spécification du sexe. Pourtant, la nature contrastive du lien apparaît comme faible, car :

1. ce qui domine avant tout dans le sens de chacune des lexies est leur nature relationnelle '[X être] fils/fille de Y' ;
2. cette relation reste identique dans les deux cas.

À l'inverse, plus des lexies sémantiquement relationnelles tendront à désigner des relations singularisées, plus l'opposition contrastive sera manifeste, comme par exemple dans le cas de MÈRE \sim PÈRE. La mère de X l'a mis au monde, alors que son père l'a engendré ; la différence est significative et a d'ailleurs des conséquences directes sur les liens paradigmatiques et syntagmatiques qui sont issus des deux lexies en question.

Remerciements

Merci à Delphine Beausery, Pascale Bernard, Anaïs Ferté, Igor Mel'čuk et Dorota Sikora pour leurs commentaires.

Bibliographie

Nabil Gader, Veronika Lux-Pogodalla et Alain Polguère. Hand-Crafting a Lexical Network With a Knowledge-Based Graph Editor. Dans *Proceedings of the Third Workshop on Cognitive Aspects of the Lexicon (CogALex III)*, pages 109–125, Mumbai, 15 décembre 2012. The COLING 2012 Organizing Committee. URL <http://www.aclweb.org/anthology/W12-5109>.

Anne-Laure Jousse. *Modèle de structuration des relations lexicales fondé sur le formalisme des fonctions lexicales*. Thèse de doctorat, Département de linguistique et traduction, Université de Montréal & UFR de linguistique, Université Paris Diderot (Paris 7), Montréal & Paris, mai 2010.

11. Pour un exemple de ce qu'il ne faut pas faire, on pourra se reporter à l'entrée « **ami, ie** » du *Petit Robert*, où le lecteur peine à distinguer ce qui relève de la description de lexies d'un vocable AMI et ce qui relève spécifiquement d'un vocable AMIE.

- Veronika Lux-Pogodalla et Alain Polguère. Construction of a French Lexical Network: Methodological Issues. Dans *Proceedings of the First International Workshop on Lexical Resources, WoLeR 2011. An ESSLLI 2011 Workshop*, pages 54–61, Ljubljana, Slovenia, August 1–5 2011.
- Igor Mel'čuk. Lexical Functions: A Tool for the Description of Lexical Relations in the Lexicon. Dans Leo Wanner, dir., *Lexical Functions in Lexicography and Natural Language Processing*, volume 31 de *Language Companion Series*, pages 37–102. John Benjamins, Amsterdam/Philadelphia, 1996.
- Igor Mel'čuk. Un FOU/une FOLLE : un lexème ou deux? *Bulag (Bulletin de Linguistique Appliquée et Générale)*, pages 95–106, 2000.
- Igor Mel'čuk. The Notion of Inflection and the Expression of Nominal Gender in Spanish. *Studies in Language*, à paraître.
- Igor Mel'čuk, André Clas et Alain Polguère. *Introduction à la lexicologie explicative et combinatoire*. 256 pages. Duculot, Paris/Louvain-la-Neuve, 1995.
- Igor Mel'čuk et Alain Polguère. Dérivations sémantiques et collocations dans le DiCo/LAF. *Langue française*, 150:66–83, 2006. Numéro spécial sur la collocation *Collocations, corpus, dictionnaires*, sous la direction de Peter Blumenthal et Franz J. Hausmann.
- Jasmina Milićević et Alain Polguère. Ambivalence sémantique des noms de communication langagière du français. Dans F. Neveu, V. Muni Toke, J. Durand, T. Klingler, L. Mondada et S. Prévost, dir., *Actes de la section « Lexique et morphologie » du 2^e Congrès Mondial de Linguistique Française (CMLF'10)*, pages 1029–1050, La Nouvelle-Orléans, 12–15 juillet 2010. Institut de Linguistique Française (ILF).
- María Moliner. *Diccionario de uso del español*. Gredos, Madrid, 2000. primera reimpression, edición abreviada por la Editorial Gredos.
- Alain Polguère. Towards a theoretically-motivated general public dictionary of semantic derivations and collocations for French. Dans *Proceedings of EURALEX'2000*, pages 517–527, Stuttgart, 2000.
- Alain Polguère. Lexical function standardness. Dans Leo Wanner, dir., *Selected Lexical and Grammatical Issues in the Meaning-Text Theory. In Honour of Igor Mel'čuk*, volume 84 de *Language Companion Series*, pages 43–95. John Benjamins, Amsterdam/Philadelphie, 2007.
- Anna Wierzbicka. *Semantic Primitives*. Linguistische Forschungen 22. Athenäum, Frankfurt, 1972.