

## ProLMF 3.1

### Sommaire

<b>1</b>	<b>Quoi de neuf</b> .....	<b>1</b>
<b>2</b>	<b>Introduction</b> .....	<b>1</b>
<b>3</b>	<b>Contributions et remerciements</b> .....	<b>2</b>
<b>4</b>	<b>Bibliographie</b> .....	<b>3</b>
<b>5</b>	<b>La modélisation des noms propres dans Prolexbase</b> .....	<b>3</b>
5.1	Les concepts.....	3
5.1.1	Le pivot.....	3
5.1.2	Le prolexème.....	4
5.1.3	Les alias et dérivés.....	5
5.1.4	Les instances.....	5
5.2	Les relations.....	6
5.2.1	La relation de synonymie.....	6
5.2.2	La relation de méronymie.....	7
5.2.3	La relation d'accessibilité.....	7
5.2.4	Les relations d'expansion.....	8
5.2.5	L'éponymie.....	8
5.3	Les relations d'hyponymie.....	8
5.3.1	La typologie des noms propres de Prolexbase.....	8
5.3.2	La typologie d'existence de Prolexbase.....	9
<b>6</b>	<b>Couverture linguistique de Prolexbase 3.1</b> .....	<b>9</b>
<b>7</b>	<b>ProLMF : la version LMF de Prolexbase</b> .....	<b>10</b>
7.1	Les catégories de données.....	10
7.2	Les classes LMF utilisées.....	10
7.3	Un extrait de ProLMF.....	13
<b>8</b>	<b>Participants au projet Prolex</b> .....	<b>14</b>

## 1 Quoi de neuf

La version 3.1 contient une révision complète de la partie serbe et son organisation suivant les préconisations de la norme LMF. Les balises *WordForm* contiennent maintenant systématiquement des balises *FormRepresentation*, ce qui autorise d'y placer des variantes, des alphabets différents et des dialectes. Dans la version 2, il y avait deux lexiques serbes, un en alphabet cyrillique et l'autre en alphabet latin ; dans la version 3, il n'y en a plus qu'un.

La couverture linguistique de Prolexbase 3.1 est présentée section 6.

La version 3.2 en cours de réalisation comportera des entrées en arabe.

Les auteurs remercient vivement ceux qui, après téléchargement, ont pris la peine de nous signaler des erreurs. Tout particulièrement Pascal Malaise, Karèn Fort et Benoît Sagot.

## 2 Introduction

Prolexbase est un lexique relationnel multilingue de noms propres. La version 3.1 comprend essentiellement du français, de l'anglais, du polonais et du serbe, mais aussi quelques traductions (sous forme de lemmes uniquement) en d'autres langues : allemand, coréen, italien, néerlandais et portugais.

La modélisation du domaine des noms propres définie dans le projet Prolex repose sur deux concepts centraux : le *pivot* et le *prolexème*. Le pivot ne représente pas le référent, mais un point de vue sur ce référent. Il possède dans chaque langue un concept spécifique, le prolexème, qui est une famille structurée de lexèmes. Autour d'eux, sont définis d'autres concepts et des relations (synonymie, méronymie, accessibilité, éponymie, etc.). Chaque pivot est en relation d'hyponymie avec un type et un paradigme d'existence.

Il n'est pas évident de définir la notion de nom propre. La plupart des définitions insistent sur le caractère unique de son référent et sur une sémantique et une syntaxe qui lui est propre. Nous avons choisi d'adopter le point de vue de (Jonasson, 1994) qui propose une définition plus large incluant ce qu'elle appelle les noms propres purs (noms de personne et noms de lieu) et les noms propres descriptifs qui résultent souvent de la composition d'un nom propre avec une expansion (Tour Eiffel, musée Rodin, etc.). Un nom propre descriptif peut être considéré comme une expression définie figée ou en cours de figement (Jardin des Plantes, Médecins sans frontières, etc.). Cette définition est assez proche de celle utilisée dans le domaine du Tal depuis la conférence MUC6.

Alors que la version 1.0 se présentait sous une forme plutôt "terminologique", puisque les entrées correspondaient aux pivots, les versions suivantes ont adopté les recommandations de la norme [LMF \(ISO-24613:2008\)](#) et se présente donc sous une forme dictionnaire où les entrées sont des lemmes.

Prolexbase 3.1 est en accès libre sous acceptation de la licence LGPL-LR. La maintenance et la mise à jour du lexique sont assurées par le [Laboratoire d'informatique](#) (LI) de l'[université François Rabelais Tours](#), l'hébergement, par le CNRTL (ATILF).

### 3 Contributions et remerciements

Prolexbase est [un projet Tal](#) du LI, en collaboration avec :

- le groupe de recherche *Langues et Représentation* (L&R) de l'université François-Rabelais, aujourd'hui Laboratoire ligérien de linguistique, UMR 7270 ;
- l'université de Belgrade ;
- l'académie des sciences de Varsovie.

La conception de la base est principalement due à Denis Maurel (Lifat), Mickaël Tran (thèse au Lifat), Thierry Grass (L&R), Duško Vitas (université de Belgrade), Agata Savary (Lifat) et Béatrice Bouchou (Lifat) ; son implantation à Denis Maurel, Mickaël Tran, Marie Ndiaye et Coralie Villes ; son installation sur le site du CNRTL à Étienne Petitjean (Atilf). La version 1.1 a été conçue par Béatrice Bouchou et Denis Maurel, avec l'aide d'Estelle Leton pour son implantation. La version 2.2 a été largement augmentée par Małgorzata Spędzia-Baron, en particulier avec l'ajout du polonais et de l'anglais, ajout accompagné du français lorsque le prolexème correspondant n'existait pas. La version 3.1 a été réalisée en collaboration pour la partie serbe avec Duško Vitas et Cvetana Krstev (université de Belgrade), suite à une invitation d'un mois financée par l'université de Tours.

Prolexbase a bénéficié de l'aide et des contributions de tous les participants au projet Prolex (voir section 8). Ce projet a reçu le soutien :

- de l'action [Technolangue](#) du Ministère de l'Industrie (2003-2005) ;
- du programme d'action intégré Egide [Pavle-Savic](#) du Ministère des Affaires étrangères (2004-2005) ;
- du projet Feder Région Centre [Entités nommées et nommables](#) (2009-2010) ;

- du projet ERDF [Nekst](#) (2009-2014) ;
- du projet européen (CIP ICT-PSP) [Cesar](#) (2011-2013).

## 4 Bibliographie

Principalement, pour la version 1.1 :

Tran M. (2006). [Prolexbase. Un dictionnaire relationnel multilingue de noms propres : conception implantation et gestion en ligne](#). Thèse de doctorat d'informatique. Université François Rabelais Tours.

Tran M., Maurel D. (2006). [Prolexbase : Un dictionnaire relationnel multilingue de noms propres](#). Traitement automatique des langues. 47(3):115-139.

Bouchou B., Maurel D. (2008). [Prolexbase et LMF : vers un standard pour les ressources lexicales sur les noms propres](#). Traitement automatique des langues. 49(1):61-88.

Pour la version 1.2 :

Bouchou, B., Maurel, D. (2013). Prolmf, a Multilingual Dictionary of Proper Names and their Relations. Chapter in Gil Francopoulo (editor), LMF: Lexical Markup Framework, theory and practice. Hermes-science, 67-81.

Pour la version 2.2 :

Savary A., Manicki L., Baron M. (2013). [Populating a Multilingual Ontology of Proper Names from Open Sources](#). Journal of Language Modelling, 1(2).

## 5 La modélisation des noms propres dans Prolexbase

L'ensemble de notre modèle se présente sous la forme d'une arborescence qui peut se décomposer en deux niveaux distincts :

- Le niveau qui ne dépend pas de la langue : le pivot et certaines relations ; la typologie des noms propres et la typologie d'existence.
- Le niveau qui dépend de la langue : le prolexème, les alias, les dérivés et certaines relations ; les instances (ensemble des formes fléchies d'une langue).

Après avoir présenté les concepts clés de notre modélisation des noms propres, nous décrirons les relations existant dans notre dictionnaire, puis nos typologies.

### 5.1 Les concepts

#### 5.1.1 Le pivot

Pour une langue donnée, des noms propres totalement différents sur le plan graphique peuvent renvoyer à un même et unique référent et ce phénomène se retrouve généralement d'une langue à l'autre. Par exemple, les noms propres *Jean-Paul II* et *Karol Jozef Wojtyla* en français correspondent tous les deux à un certain point de vue sur un même et unique référent et il en est de même en anglais (*John Paul II* et *Karol Jozef Wojtyla*), en italien (*Giovanni Paolo II* et *Karol Jozef Wojtyla*), etc.

Nous définissons donc le pivot non pas comme le référent, mais, plutôt, comme un certain point de vue sur celui-ci. Il sera représenté dans la base par un numéro identifiant unique. Ainsi les noms propres *Pologne* en français, *Polonia* en espagnol, *Polen* en allemand, etc. seront associés à un même pivot, tandis que les noms propres *République de Pologne* en

français, *República Polaca* en espagnol, *Republik Polen* en allemand, etc. seront associés à un autre pivot. Ces deux pivots seront en relation de synonymie.

Pour définir ces différents points de vue, nous nous sommes basés sur un marquage diasystématique, provenant des travaux sur la métalexigraphie d'Eugenio Coseriu qui propose un diasystème basé essentiellement sur quatre variétés distinctes : diachronique (variété dans le temps), diaphasique (variété concernant les finalités de l'emploi), diatopique (variété dans l'espace) et diastratique (variété relative à la stratification socio-culturelle).

### 5.1.2 Le prolexème

Dans notre modèle, le prolexème correspond à une projection du pivot dans une langue donnée. Chaque prolexème d'une langue donnée sera donc relié à un seul et unique pivot. C'est en se basant sur cette relation que l'on va pouvoir traduire les prolexèmes d'une langue vers une autre. Le concept de prolexème peut aussi se définir comme une classe d'équivalence.

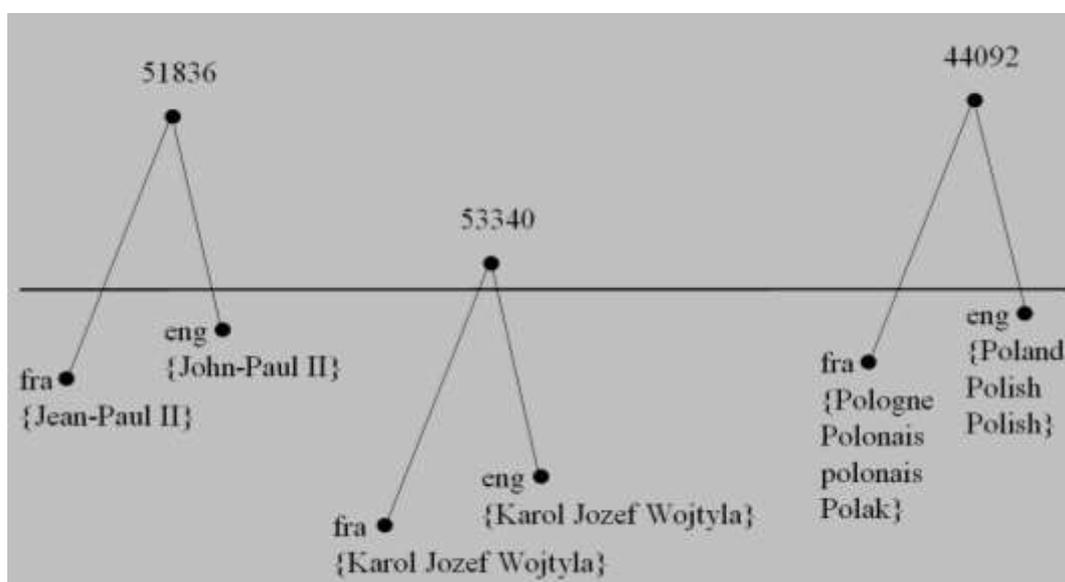


Figure 1 : Pivots et prolexèmes

Pour simplifier, nous considérons aussi le prolexème comme le lemme associé aux différentes formes d'un nom propre qui apparaissent dans les différents textes d'une langue donnée. Il peut ainsi être considéré comme la forme vedette d'un ensemble de dérivés et d'alias.

Les noms propres polysèmes, qui sont classés sous des catégories différentes, seront reliés à des prolexèmes différents. Par exemple, *Verdun* est à la fois connu comme étant une célèbre bataille durant la Première Guerre Mondiale, comme un traité entre les trois fils de l'empereur Louis le Pieux pour partager son Empire et, enfin, comme le chef-lieu de la Meuse. Pour ce cas-là, nous serons amenés à créer trois prolexèmes différents. Par contre, dans le cas de toponymes correspondant à la fois à un lieu et une entité administrative (comme par exemple *Paris* qui est à la fois une ville et un département), nous avons décidé de ne pas dupliquer les prolexèmes pour éviter l'abondance d'homographes. Cette information sera rajoutée au niveau des expansions classifiantes du prolexème.

Les noms propres homographes seront aussi associés à des prolexèmes différents. En recherchant le nom propre *Sydney* dans un dictionnaire, on trouvera deux entrées distinctes : une qui correspondra à une ville en Australie et l'autre à une ville située au Canada. Il est à noter que l'homonymie dépend de la langue. Par exemple, en anglais, le nom propre *London*

correspond au chef-lieu de l'île Kiritimati ou à la capitale de l'Angleterre, ce qui n'est pas le cas en français à cause de l'existence d'un exonyme (*Londres*).

### 5.1.3 Les alias et dérivés

Nous définissons les alias comme des synonymes qui dépendent de la langue. Nous avons regroupé dans le terme d'alias, d'une part, des synonymes exacts (les variantes d'écriture - caractères, abréviations, acronymes et sigles, transcriptions-, les variantes orthographiques) et, d'autre part, des quasi-synonymes, diatopiques ou diastratiques.

Par exemple, les noms *Nations Unies*, *Onusien*, *ONU* auront *Organisation des Nations Unies* comme prolexème pour la langue française. Les noms *United Nations* et *UNO* auront pour prolexème *United Nations Organization* pour la langue anglaise. Le prolexème français *Organisation des Nations Unies* et le prolexème anglais *United Nations Organization* seront reliés à un même pivot (48226).

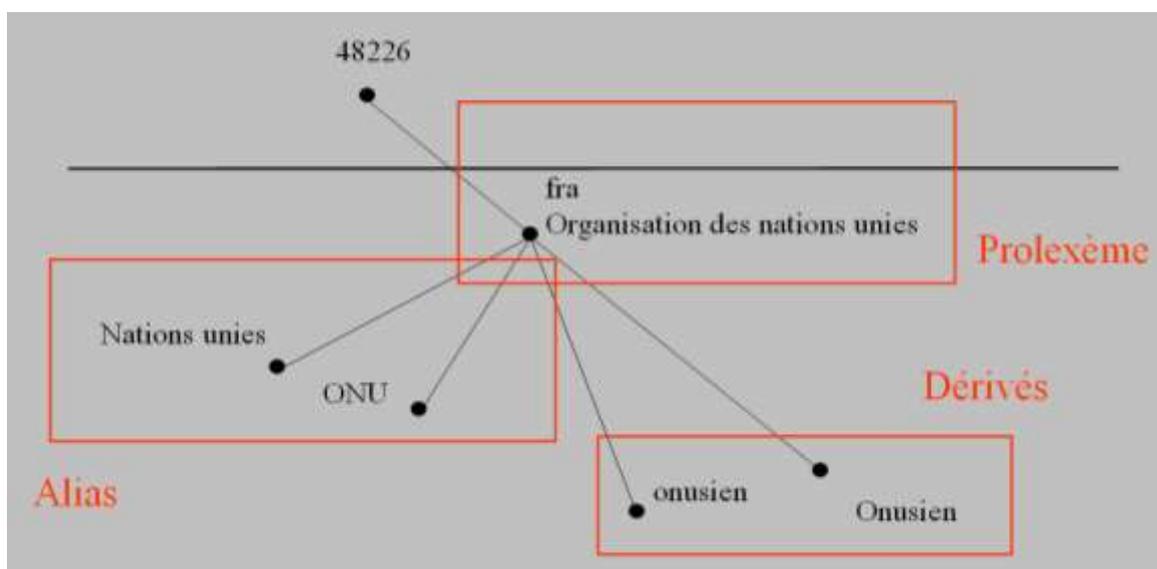


Figure 2 : Alias et dérivés

Les dérivés présents dans Prolexbase sont des dérivés morphosémantiquement liés aux prolexèmes. Ainsi *polonais* y figure (*le président de Pologne* versus *le président polonais*), mais pas *pasteuriser* (qui est lexicalisé dans un sens très précis et qui ne peut se paraphraser à partir du prolexème : *\*faire comme Pasteur faisait*).

Le classement des alias et des dérivés dans la partie qui dépend de la langue s'explique notamment par la raison que la créativité lexicale est propre à chaque langue. Une variante d'écriture ou un dérivé existant dans une langue  $L_1$  peut être totalement absent dans une langue  $L_2$ . Par exemple, le dérivé *Tourangeau* se traduira en anglais par *inhabitant of Tours*.

### 5.1.4 Les instances

Les instances sont toutes les formes fléchies que l'on peut obtenir à partir d'un nom propre, d'un de ses alias ou d'un de ses dérivés.

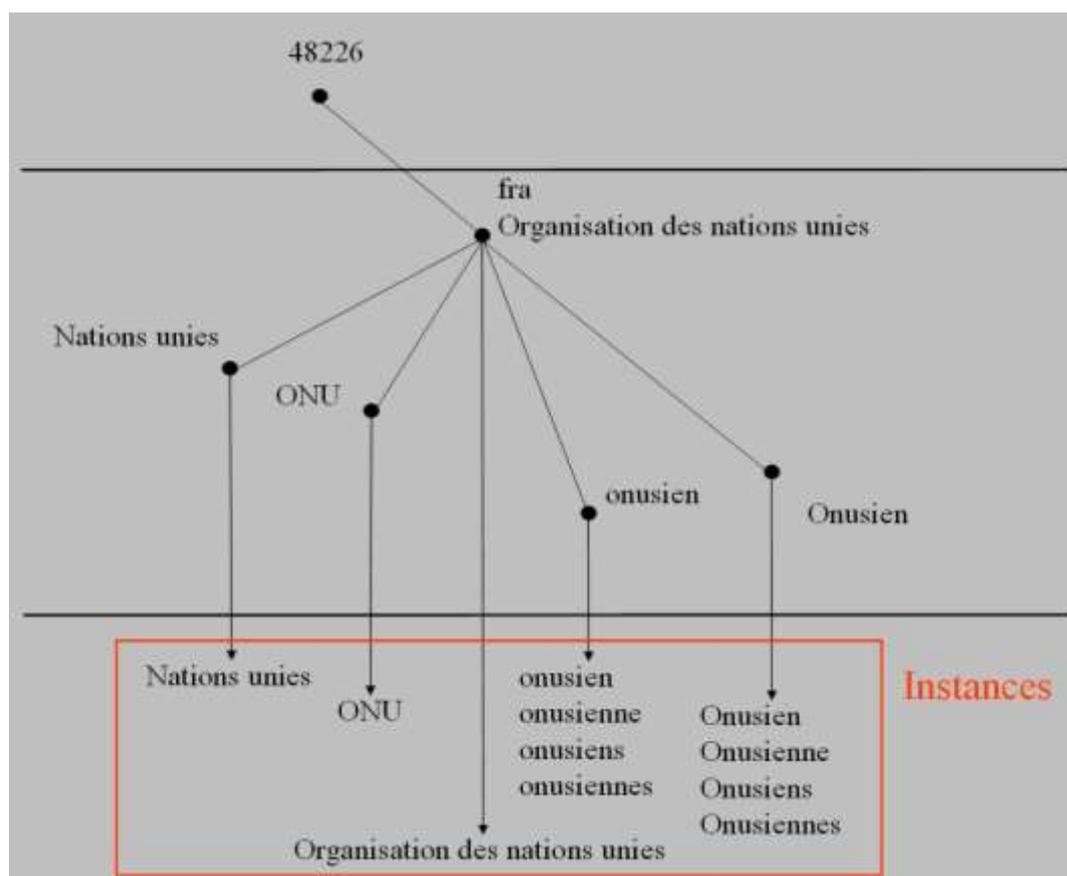


Figure 3 : Instances

## 5.2 Les relations

Dans le cas des relations qui ne dépendent pas de la langue, nous avons retenu pour Prolexbase trois relations paradigmatiques (méronymie, synonymie, hyperonymie) et une relation syntagmatique (accessibilité). La relation d'hyperonymie correspond, dans Prolexbase, à une typologie des noms propres et à un paradigme d'existence que nous présenterons plus loin.

Pour le français uniquement, dans le cas des relations qui dépendent de la langue, nous avons retenu quatre relations syntagmatiques : une collocation figée (l'éponymie) et trois collocations libres :

- La présence éventuelle d'un déterminant (*la France*) et le choix de la préposition locative pour les noms de pays (*en France*).
- l'expansion classifiante (*la ville de Paris*)
- l'expansion d'accessibilité (*Paris, la capitale de la France*)

### 5.2.1 La relation de synonymie

Dans une synonymie, l'un des termes est souvent préférable à l'autre. On appellera le premier la forme canonique et l'autre la forme synonyme. Cette forme canonique en général correspond à la forme la plus connue. Nous avons considéré la variation diatopique comme un alias (voir plus haut). Il reste donc, au niveau indépendant de la langue, les trois variations : diachronique (*Zaire* et *République démocratique du Congo*), diastratique (*Molière* et *Jean-Baptiste Poquelin*) et diaphasique (*Paris* et *Ville Lumière*).

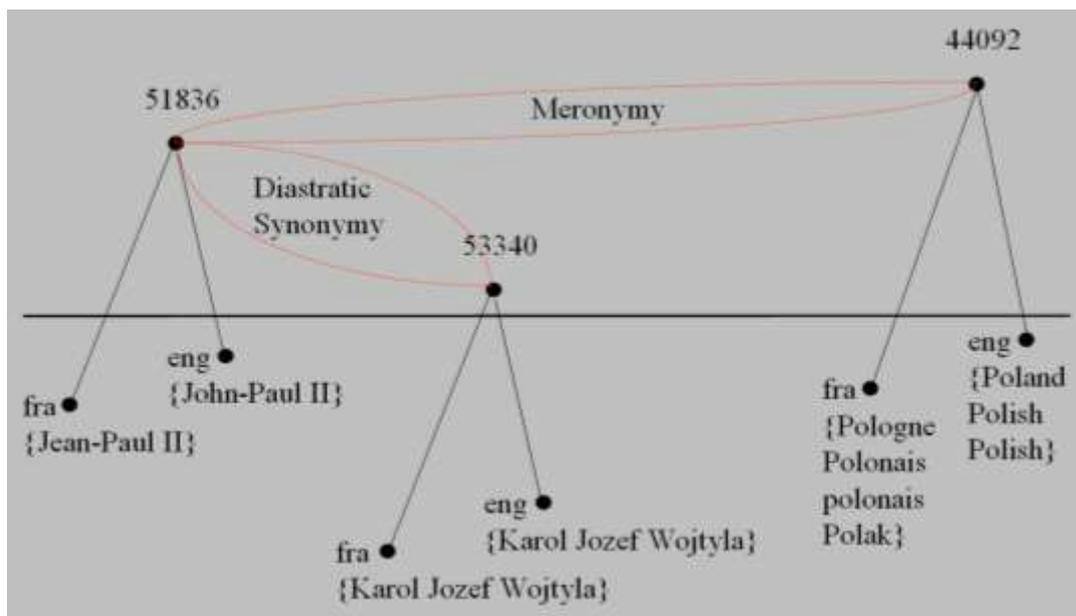


Figure 4 : Relations qui ne dépendent pas de la langue

### 5.2.2 La relation de méronymie

La relation de méronymie permet d'établir une hiérarchisation sur plusieurs niveaux entre les éléments contenant (holonymes) et les éléments contenus (méronymes). Lorsque deux unités lexicales A et B sont en relation de méronymie, on dit que A est un méronyme de B (et que B est un holonyme de A) si A est une partie de B.

Nous entendons cette relation au sens large, par exemple une célébrité et son pays, un roi et sa dynastie, un personnage et une œuvre, des toponymes entre eux, des événements entre eux, etc.

### 5.2.3 La relation d'accessibilité

Un nom propre dans un dictionnaire n'est pas associé à une définition classique, mais à une description encyclopédique, faite avec d'autres noms propres sur lesquels se base son accessibilité. Cette relation pourrait en fait correspondre à une multitude de relations, mais cela risquerait d'être coûteux et de nuire à la lisibilité du modèle, d'autant que certaines expansions existent dans une langue et sont absentes dans d'autres langues. Par exemple, en France, nous faisons la distinction entre un chef-lieu, une préfecture, une capitale, etc. Nous avons donc décidé de les regrouper dans une seule et unique relation, à laquelle nous avons ajouté des repérages généraux. Les informations spécifiques sont conservées dans la relation d'expansion classifiante (voir ci-dessous). Cette relation d'accessibilité n'est pas une modélisation idéale, mais correspond à une solution économique, suffisante pour la plupart des applications de Tal.

Donnons quelques exemples de repérages qui peuvent apparaître dans le cadre d'une relation d'accessibilité et leur contexte en français :

- Parent : les personnes et les membres de leur famille. *Joseph* est la mère de *Jésus*, *Henri II* est le fils de *François I<sup>er</sup>*, etc.
- Créateur : les auteurs et les œuvres. *Molière* est l'auteur de *Tartuffe*, etc.
- Capitale : les toponymes et leurs capitales. *Tours* est le chef-lieu de *l'Indre et Loire*, *Bangkok* est la capitale de la *Thaïlande*, etc.

- Dirigeant politique : les hommes politiques et les pays. *Jacques Chirac* est un homme politique *français*, etc.
- Siège : les entreprises, associations ou organisations et le toponyme correspondant au siège social. *Peugeot* est une firme *sochaliennne*, etc.
- Etc.

#### 5.2.4 Les relations d'expansion

Ces relations associent à un prolexème, ou à deux prolexèmes dont les pivots sont en relation d'accessibilité, une expansion qui peut apparaître soit à sa gauche, soit à sa droite. Toutes les expansions qui existent dans une langue ne se retrouvent pas forcément dans une autre langue. Si l'expansion d'un nom propre est omise dans un texte, il est parfois nécessaire de la rétablir lors de la traduction de celui-ci, afin d'apporter un complément d'information au lecteur. Ainsi, le nom propre *la Loire* devient en anglais *the Loire River*. Comme il a été dit ci-dessus, certaines expansions complètent la notion de repérage associée à la relation d'accessibilité entre deux noms propres.

#### 5.2.5 L'éponymie

Les noms propres apparaissent parfois dans les textes sous la forme de simples substantifs. Cette possibilité existe dans un grand nombre de langues, mais, pour un nom propre donné, dépend de la langue considérée. La relation d'éponymie, non renseignée à ce jour, est donc la relation entre un nom propre et une forme lexicalisée, soit dans le vocabulaire courant (*un tartuffe*), soit dans une expression idiomatique (*tous les chemins mènent à Rome*) ou terminologique (*maladie de Parkinson*). Contrairement aux autres relations, l'objectif de la prise en compte de la relation d'éponymie est d'empêcher une reconnaissance abusive des noms propres dans des textes.

### 5.3 Les relations d'hyponymie

#### 5.3.1 La typologie des noms propres de Prolexbase

La typologie des noms propres de Prolexbase est hiérarchisée par une relation d'hyponymie qui a pour racine le concept de nom propre, pour nœuds des supertypes et pour feuilles des types.

Situés juste en dessous du concept de nom propre, les quatre premiers supertypes classent les noms propres suivant des traits syntaxico-sémantiques assez généraux : les *anthroponymes* (trait humain), les *ergonymes* (les fabrications humaines - trait inanimé), les *pragmonymes* (trait événement) et les *toponymes* (trait locatif).

Nous avons partagé le supertype anthroponyme en deux autres supertypes : les *anthroponymes individuels* et les *anthroponymes collectifs*.

Les types correspondent à une classification plus détaillée, mais volontairement limitée. Dans Prolexbase, nous avons retenu au total trente types (voir Figure 5). Comme certaines distinctions sont difficiles à réaliser et peuvent sembler arbitraires, nous avons créé deux autres supertypes : un supertype hyponyme des *anthroponymes collectifs*, les *groupements*, et un supertype hyponyme des *toponymes*, les *territoires*.

Nom propre						
Anthroponyme			Toponyme		Ergonyme	Pragmonyme
Individuel	Collectif		Astronyme Edifice Géonyme Hydronyme Ville Voie	Territoire	Objet Œuvre Pensée Produit Vaisseau	Catastrophe Fête Histoire Manifestation Météorologie
Personne Patronyme Prénom Pseudo Anthroponyme	Dynastie Ethnonyme	Groupement		Pays Région Supranational		
		Association Ensemble Entreprise Institution Organisation				

Figure 5 : Les types de Prolexbase

Dans les versions téléchargeables de Prolexbase, chaque pivot est en relation d'hyponymie avec un type. Donc seuls ceux-ci apparaissent.

### 5.3.2 La typologie d'existence de Prolexbase

L'hyponymie d'*existence* relie les pivots à l'une des trois instances :

- Historique : Etre un nom propre du domaine historique (*Mozart, le Danube, Paris*, etc.).
- Religieux : Etre un nom propre du domaine de la croyance (*Zeus, Adam, Gabriel*, etc.).
- Fictif : Etre un nom propre du domaine de la fiction (*Don Quichotte de la Manche, Eldorado*, etc.).

La distinction entre les noms propres historiques et les autres s'avère utile pour la traduction, car, dans la majorité des cas, ces derniers possèdent des traductions distinctes d'une langue à l'autre. Par exemple, *Don Quijote de la Mancha* en espagnol devient *Don Quichotte de la Manche* en français et *Don Quixote of La Mancha* en anglais.

## 6 Couverture linguistique de Prolexbase 3.1

Prolexbase 3.1 contient :

- 73 510 pivots
- 72 572 relations : 3 906 accessibilités, 67 391 méronymies et 1 275 synonymies
- 96 800 entrées lexicales en français : 79 194 noms propres et 21 299 dérivés, soit 148 692 formes
- 19 941 entrées lexicales en anglais : 19 847 noms propres et 94 dérivés, soit 19 446 formes
- 39 394 entrées lexicales en polonais : 36 312 noms propres et 3 082 dérivés, soit 153 999 formes
- 8 986 entrées lexicales en serbe : 8 546 noms propres et 440 dérivés, soit 54 041 formes

À ces lexiques s'ajoutent quelques petits lexiques d'autres langues, ne comprenant malheureusement que les lemmes et les sens : allemand, coréen, italien, néerlandais et portugais.

## **7 ProLMF : la version LMF de Prolexbase**

### **7.1 Les catégories de données**

Si les classes et leurs liens sont normalisés par LMF, il n'en est pas de même pour les attributs que l'on souhaite leur attacher. Cependant, il est recommandé de suivre autant que faire se peut la norme [ISO 12620](#) qui « spécifie les catégories de données utilisées pour l'enregistrement de l'information terminologique [...] ainsi que pour l'échange et la recherche d'information terminologique ». Cette norme est complétée par [[Francopoulo et al., 2008](#)] et par quelques ajouts de notre part. Le fichier ProLMF\_3\_1.xsd joint dans l'archive détaille les choix que nous avons faits.

### **7.2 Les classes LMF utilisées**

Ce même fichier contient le schéma XML de ProLMF. Nous utilisons d'abord le noyau obligatoire de LMF, avec la langue et, à l'intérieur d'une langue, les formes et les sens d'une entrée. Les formes correspondent aux lemmes (prolexèmes, alias et dérivés) de Prolexbase, ainsi qu'à leurs instances.

Pour les sens, Prolexbase distingue les relations qui ne dépendent pas de la langue (synonymie, méronymie et accessibilité) et les relations qui en dépendent (aliasation, dérivation morphosémantique, expansion classifiante, éponymie). Il s'agit de ne pas dupliquer dans chaque langue une même information. C'est pourquoi nous utilisons les relations de l'extension multilingue pour les premières et les relations de l'extension sémantique pour les secondes.

La Figure 6 présente les classes utilisées dans ProLMF. La partie grisée correspond à l'extension multilingue.

L'extension morphologique ne prévoit qu'un unique lemme par entrée lexicale. Les formes homographes sont donc rassemblées en une seule entrée, chaque nom propre étant distingué au niveau du sens. Nous présentons ci-dessous les entrées lexicales (Figure 7), les relations dépendantes (Figure 8) et indépendantes (Figure 9) de la langue.

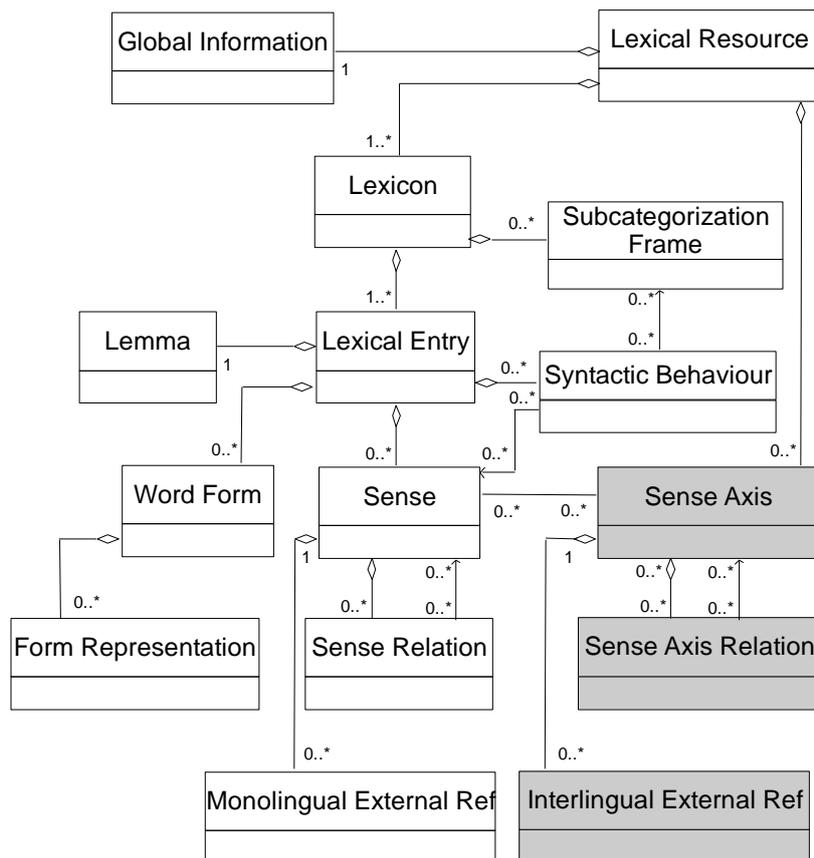


Figure 6 : Les classes utilisées par ProLMF

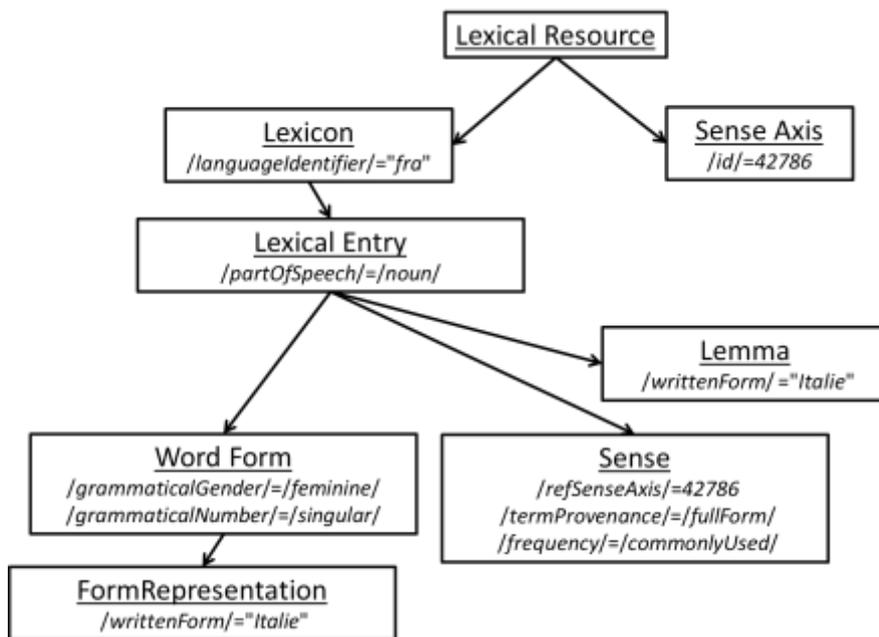


Figure 7 : Une partie de l'entrée lexicale *Italie*

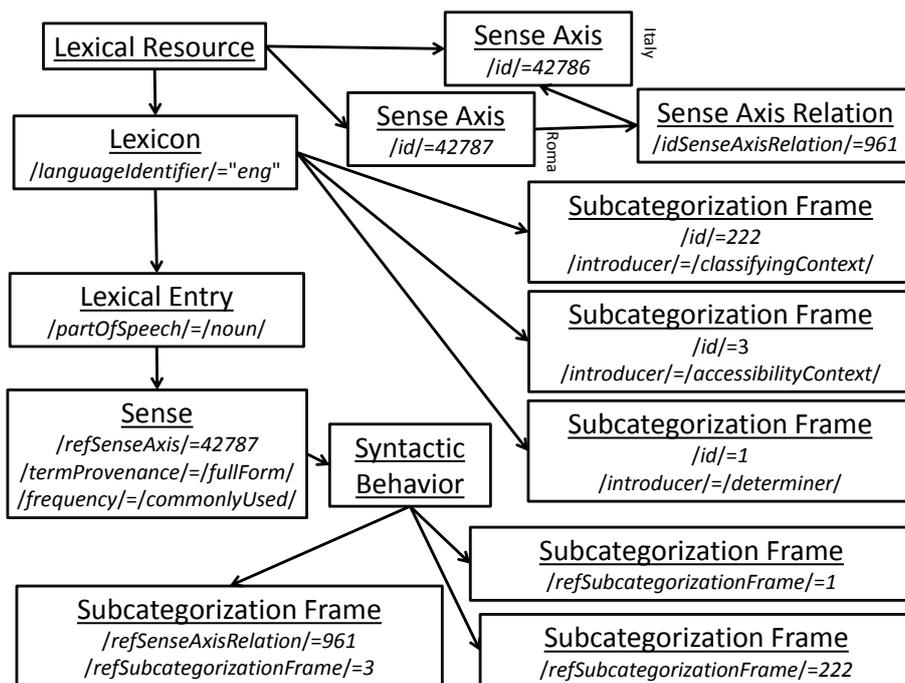


Figure 8 : Les relations dépendantes de la langue

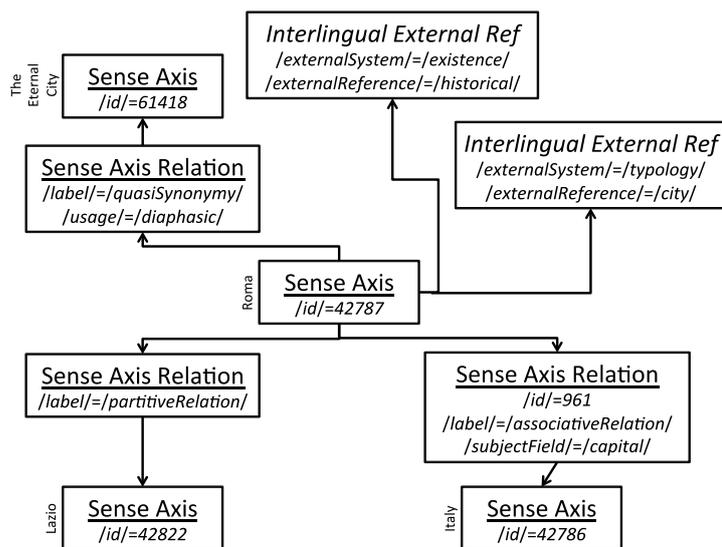


Figure 9 : Les relations indépendantes de la langue

### 7.3 Un extrait de ProLMF

L'extrait ci-dessous de ProLMF correspond aux noms propres *Italie*, *Latium* et *Rome* et au dérivé *Italien*.

```

<?xml version='1.0' encoding='UTF-8' ?>
<LexicalResource>
  <GlobalInformation languageCoding="ISO 639" scriptCoding="ISO 15924" characterCoding="UTF-8"
entrySource="Prolexbase" resourceName="Prolmf" version="1.2"/>
  <Lexicon languageIdentifier="fra" script="latn">
    <LexicalEntry partOfSpeech="noun">
      <Lemma>Italie</Lemma>
      <WordForm grammaticalGender="feminine" grammaticalNumber="singular">
        <FormRepresentation>Italie</FormRepresentation>
      </WordForm>
      <Sense idSense="P42786" refSenseAxis="42786" termProvenance="fullForm"
frequency="commonlyUsed" label="properName">
        <SyntacticBehaviour refSubcategorizationFrame="CO4"/>
        <SyntacticBehaviour refSubcategorizationFrame="CO7"/>
        <MonolingualExternalRef externalSystem="Wikipedia"
externalReference="http://fr.wikipedia.org/wiki/Italie"/>
      </Sense>
    </LexicalEntry>
    <LexicalEntry partOfSpeech="noun">
      <Lemma>Italien</Lemma>
      <WordForm grammaticalGender="masculine" grammaticalNumber="singular">
        <FormRepresentation>Italien</FormRepresentation>
      </WordForm>
      <WordForm grammaticalGender="masculine" grammaticalNumber="plural">
        <FormRepresentation>Italiens</FormRepresentation>
      </WordForm>
      <WordForm grammaticalGender="feminine" grammaticalNumber="singular">
        <FormRepresentation>Italienne</FormRepresentation>
      </WordForm>
      <WordForm grammaticalGender="feminine" grammaticalNumber="plural">
        <FormRepresentation>Italiennes</FormRepresentation>
      </WordForm>
      <Sense idSense="D19751" refSenseAxis="42786" termProvenance="relationalName"
frequency="commonlyUsed" label="derivative" refSense="P42786"/>
    </LexicalEntry>
    <LexicalEntry partOfSpeech="noun">
      <Lemma>Latium</Lemma>
      <WordForm grammaticalGender="masculine" grammaticalNumber="singular">
        <FormRepresentation>Latium</FormRepresentation>
      </WordForm>
      <Sense idSense="P42822" refSenseAxis="42822" termProvenance="fullForm"
frequency="infrequentlyUsed" label="properName">
        <SyntacticBehaviour refSubcategorizationFrame="CC179"/>
        <SyntacticBehaviour refSubcategorizationFrame="CO2"/>
        <MonolingualExternalRef externalSystem="Wikipedia"
externalReference="http://fr.wikipedia.org/wiki/Latium"/>
      </Sense>
    </LexicalEntry>
    <LexicalEntry partOfSpeech="noun">
      <Lemma>Rome</Lemma>
      <WordForm grammaticalGender="masculineFeminine" grammaticalNumber="singular">
        <FormRepresentation>Rome</FormRepresentation>
      </WordForm>
      <Sense idSense="P42787" refSenseAxis="42787" termProvenance="fullForm"
frequency="commonlyUsed" label="properName">
        <SyntacticBehaviour refSubcategorizationFrame="CC222"/>
        <SyntacticBehaviour refSenseAxisRelation="961"
refSubcategorizationFrame="AC3"/>
        <SyntacticBehaviour refSenseAxisRelation="961"
refSubcategorizationFrame="AC52"/>
        <SyntacticBehaviour refSubcategorizationFrame="CO1"/>
        <MonolingualExternalRef externalSystem="Wikipedia"
externalReference="http://fr.wikipedia.org/wiki/Rome"/>
      </Sense>
    </LexicalEntry>
  </Lexicon>
</LexicalResource>

```

.../...

```

<SubcategorizationFrame id="CO1" introducer="determiner">zero</SubcategorizationFrame>
<SubcategorizationFrame id="CO2" introducer="determiner">le</SubcategorizationFrame>
<SubcategorizationFrame id="CO3" introducer="determiner">la</SubcategorizationFrame>
<SubcategorizationFrame id="CO4" introducer="determiner">'</SubcategorizationFrame>
<SubcategorizationFrame id="CO7" introducer="locativePreposition">en</SubcategorizationFrame>
<SubcategorizationFrame id="CC179" introducer="classifyingContext" grammaticalGender="feminine"
grammaticalNumber="singular">la région de </SubcategorizationFrame>
<SubcategorizationFrame id="CC222" introducer="classifyingContext" grammaticalGender="feminine"
grammaticalNumber="singular">la ville de </SubcategorizationFrame>
<SubcategorizationFrame id="AC3" introducer="accessibilityContext" grammaticalGender="feminine"
grammaticalNumber="singular">la capitale de l'</SubcategorizationFrame>
<SubcategorizationFrame id="AC52" introducer="accessibilityContext" grammaticalGender="feminine"
grammaticalNumber="singular">le chef-lieu du </SubcategorizationFrame>
</Lexicon>
<SenseAxis id="42786">
<InterlingualExternalRef externalSystem="typology" externalReference="country"/>
<InterlingualExternalRef externalSystem="existence" externalReference="historical"/>
<SenseAxisRelation label="partitiveRelation" refSenseAxis="42787"/>
<SenseAxisRelation label="partitiveRelation" refSenseAxis="42822"/>
</SenseAxis>
<SenseAxis id="42787">
<InterlingualExternalRef externalSystem="typology" externalReference="city"/>
<InterlingualExternalRef externalSystem="existence" externalReference="historical"/>
<SenseAxisRelation id="961" label="associativeRelation" refSenseAxis="42786" subjectField="capital"/>
<SenseAxisRelation id="962" label="associativeRelation" refSenseAxis="42822" subjectField="capital"/>
</SenseAxis>
<SenseAxis id="42822">
<InterlingualExternalRef externalSystem="typology" externalReference="region"/>
<InterlingualExternalRef externalSystem="existence" externalReference="historical"/>
<SenseAxisRelation label="partitiveRelation" refSenseAxis="42787"/>
</SenseAxis>
</LexicalResource>

```

## 8 Participants au projet Prolex

Merci à tous ceux qui, de près ou de loin, ont participé au projet Prolex :

Claire AGAFONOV	Thierry GRASS	Odile PITON
Jean-Yves ANTOINE	Assia KHEMSI	Gilles RICHARD
Claude BELLEIL	Cvetana KRSTEV	Adeline ROISSÉ
Fatma BEN MESMIA	Mélanie LEBRUN	Nathalie ROSSI
Samir BESSAIH	Émeline LECUIT	Agata SAVARY
Béatrice BOUCHOU	Benoît LEDUC	Maureen SCHEIDNES
Géraud BUXEROLLES	Hye-Kyung LEE	Elsa SKLAVOUNOU
Julien CARCENAC	Estelle LETON	Hamitouche SMAILI
Christophe DAUDIN	Denis MAUREL	Małgorzata SPĘDZIA-BARON
Elmar EGGERT	Antoine MICAUD	Catherine TAIEB
Mouna ELASHTER	Ionas MICHAÏLIDIS	Mickaël TRAN
Nathalie FRIBURGER	Hassina MOULOUDI	Coralie VILLES
Nathalie GARRIC	Valérie MOUNIER	Duško VITAS
Marie-Aimée GAZEAU	Marie NDIAYE	
Manaz GHORAISHI	Ourania PAPADIMA	