

DOROTA SIKORA
ATILF – CNRS (UMR 7118)
dorota.sikora@atilf.fr

DÉFINIR LE SENS DANS UN RÉSEAU LEXICAL*

Abstract. Dorota Sikora, *Définir le sens dans un réseau lexical* [How to define a word in a lexical network?], *Studia Romanica Posnaniensia*, Adam Mickiewicz University Press, Poznań, vol. XXXIX/3: 2012, pp. 63-79. ISBN 978-83-232-2469-3. ISSN 0137-2475. eISSN 2084-4158.

As underlined in Wierzbicka (1996: 239), a « linguistic definition is a scientific hypothesis about the concept encoded in a given word ». The present paper focuses on the problem of what such hypothesis is about. Lexicon of any language is structured with a set of connections relating lexical items. That is to say that speakers of any language need to master not only separate words, but rather a cluster of connected senses. If so, then a lexical definition should lay out the concept the word lexicalizes in the way that gives access to the whole grasp of its conceptual, paradigmatic and syntagmatic relations. We discuss the topic of definitional structures developed in the context of the French Lexical Network (Réseau Lexical du Français).

Key words : lexical definition, Explanatory and Combinatorial Lexicology, lexical resources

INTRODUCTION

Dans cet article, nous présentons les définitions lexicographiques du Réseau Lexical du Français (désormais RLF). Cette ressource lexicale se pose deux objectifs indissociablement liés. D'une part, elle fournit à l'utilisateur toutes les informations sémantiques, syntaxiques, morphologiques nécessaires pour maîtriser l'usage d'une unité lexicale. De l'autre, elle se construit en tant que modèle représentant la structuration du lexique du français sous forme d'un graphe.

Dans la section qui suit, nous présenterons brièvement le projet RLF et son cadre théorique. La section (2) nous permettra de décrire le patron général des définitions, en exposant les principes qui régissent l'élaboration des paraphrases définitoires. Dans la troisième partie, nous nous servirons de quelques exemples de définitions du RLF pour montrer les relations conceptuelles qu'elles permettent d'établir. Nous discuterons ensuite les liens paradigmatiques et syntagmatiques qu'elles décrivent (section 4). Toutes ces réflexions nous amèneront vers des conclusions exposées dans la dernière section.

* Nous tenons à remercier chaleureusement Alain Polguère pour sa relecture critique d'une première version de cet article.

Précisons dès à présent deux notions essentielles pour notre exposé. Dans les pages qui suivent, nous employons systématiquement les termes de *vocable* et d'*unité lexicale*. Le vocable est un ensemble comprenant minimalement une, et potentiellement plusieurs unités lexicales. Nous parlerons indifféremment d'*unité lexicale*, *lexie*, *acception* ou *sens*.

1. CADRE THÉORIQUE DU RLF ET PRÉSENTATION GÉNÉRALE DE LA RESSOURCE

1.1. LEXICOLOGIE EXPLICATIVE ET COMBINATOIRE

L'impulsion pour la conception du RLF vient de la réflexion menée par les linguistes de la Lexicologie Explicative et Combinatoire (désormais LEC) dans le cadre plus général de la Théorie Sens-Texte développée par Mel'čuk (pour une introduction, voir, entre autres, Mel'čuk, 1997, Milićević, 2006, Polguère 2011b). Cette théorie se caractérise par une orientation résolument lexicaliste, en plaçant l'ensemble des unités lexicales au cœur du système linguistique : « En exagérant quelque peu, on pourrait même dire que l'ensemble de lexies est la langue. En effet, une langue est constituée de lexies et de règles servant à la manipulation de celles-ci » (Mel'čuk, I., Clas, A. et A. Polguère, 1995 : 17).

Il est clair qu'une telle vision du lexique n'est pas indifférente pour le travail lexicographique. Depuis les années quatre-vingt, elle a conduit à l'élaboration de tout un ensemble de principes déterminant¹ la démarche descriptive, que l'on désigne avec Polguère (*à paraître*), sous le nom de Lexicographie Explicative et Combinatoire (LgEC).

La LEC envisage l'unité lexicale en tant que signe linguistique sous ses deux facettes saussuriennes – signifiant et signifié – en les complétant par une troisième : sa combinatoire syntaxique. La pratique de la LgEC consiste donc à analyser et à décrire ces trois aspects constitutifs, et indissociables dans le discours, du signe linguistique. Voyons à présent comment une ressource conçue dans cet esprit répond aux postulats de la LEC.

1.2. RÉSEAU LEXICAL DU FRANÇAIS

Le RLF n'est pas la première entreprise réalisée selon les préceptes de la LEC². Il suffit de rappeler les quatre volumes du Dictionnaire Explicatif et Combinatoire (Mel'čuk et *alii* 1984, 1988, 1992, 1999) qui présentent des discussions focalisées

¹ Il nous est impossible, dans le cadre du présent article, ne serait-ce que de les énumérer. Le lecteur pourra se reporter à Mel'čuk, Clas et Polguère (1995).

² Pour une discussion approfondie de ces ouvrages et ressources, voir Polguère (*à paraître*).

sur des problèmes lexico-sémantiques précis illustrées par des descriptions lexicographiques de plus de 600 vocables. Il s'agit néanmoins d'ouvrages à destination de linguistes plutôt que d'un utilisateur non spécialiste. *Le lexique actif du français* (Mel'čuk et Polguère 2007) s'adresse, au contraire, à un public plus large non initié. Il se conjugue également en version de base lexicale consultable en ligne³.

Le RLF⁴, développé actuellement au Laboratoire ATILF CNRS (UMR 7118) sous la direction scientifique d'Alain Polguère, est une base de données lexicales destinée à assurer, au terme du projet, une large couverture du lexique français. Sa première étape de trois ans consiste à proposer une description du vocabulaire de base (McCarthy, 1999), c'est-à-dire d'environ 3700 vocables dont les lexies se caractérisent par un potentiel communicatif le plus élevé dans le lexique français. Néanmoins, cette nomenclature doit rapidement s'accroître selon des règles bien déterminées (Polguère et Sikora, à paraître).

Les données lexicales du RLF permettent de traiter toutes les facettes du signe linguistique. En cas de polysémie (c'est-à-dire dans la plupart des cas), toutes les acceptions du vocable sont clairement distinguées et décrites séparément. Chaque unité lexicale (lexie-vedette) est identifiée en termes de ses caractéristiques grammaticales. Une définition analytique (voir la section suivante), les liens lexicaux (*fonctions lexicales*) qu'elle contrôle constituent la partie sémantiquement essentielle de l'article. Viennent ensuite des exemples tirés de Frantext, d'un corpus journalistique de presse régionale et d'un autre comprenant des tranches du Web. Enfin, la rubrique intitulée « Nota bene » permet de fournir d'autres informations, par exemple les connotations véhiculées par l'unité lexicale.

On imagine facilement la taille que représenterait une telle entrée dans un dictionnaire traditionnel. Par ailleurs, l'exploitation de toutes ces données saisies sous forme d'un texte serait extrêmement lourde pour les utilisateurs de la ressource. Il est donc primordial de préciser que le RLF est un dictionnaire virtuel (Selva, Verlinde et Binon, 2003). Les données sont atomisées et stockées dans des tables d'une base de données de type SQL via l'éditeur MVS_Dicet⁵. Cette structure de données permet d'envisager la génération automatique des dictionnaires de différents types. La figure 1 offre une vue d'article pour une entrée dans le RLF.

L'objectif que se pose le RLF va pourtant au-delà de création de ressources dictionnaires linéaires de par leur nature. Il s'agit de créer un modèle du lexique du français aussi fidèle que possible à l'« original ». Conformément à son nom, le RLF représente le réseau que constitue le lexique, en l'occurrence celui du français. Cela signifie que les données saisies dans la base de données doivent permettre de créer un graphe connexe d'unités lexicales.

³ <http://olst.ling.umontreal.ca/dicopop/>

⁴ Pour une présentation du projet, voir : <http://www.atilf.fr/spip.php?article908>

⁵ L'éditeur MVS_Dicet est développé par Nabil Gader de la société Micro Video Soft, partenaire du Laboratoire ATILF dans le projet RELIEF/RLF.

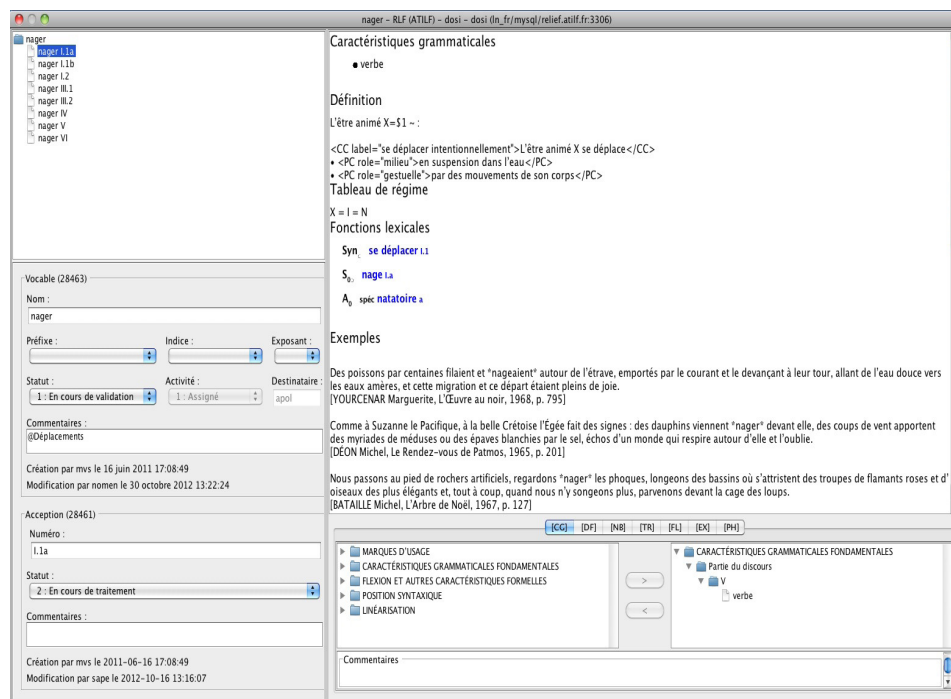


Fig. 1. Vue d'articles des unités lexicales du vocable NAGER dans le RLF

Dans les paragraphes qui suivent, nous nous concentrerons sur le volet définitionnel du RLF. Nous expliquerons comment se construisent les paraphrases définitoires dans un réseau lexico-sémantique.

2. STRUCTURE DES DÉFINITIONS DANS LE RLF

2.1. STRUCTURE DU *DEFINIENDUM*

À l'exception des ressources lexicographiques réalisées dans le cadre de la LEC que nous venons de mentionner, les dictionnaires traditionnels, indépendamment de leur mode d'existence électronique ou papier, ne diffèrent guère quant à l'objet des définitions qu'ils proposent. Le *definiendum* équivaut habituellement à la forme du vocable accompagnée d'indications grammaticales.

Conformément aux principes de la LEC, pour un lexicographe du RLF, le premier pas vers une définition consiste à dégager la forme propositionnelle propre à chaque sens à paraphraser. La forme propositionnelle est une formule explicitant sous forme de variables X, Y, Z, etc., tous les arguments sémantiques d'un prédicat. Ainsi, le verbe *NAGER* des exemples (1) et (2), tirés de Frantext, se voit associer deux formes propositionnelles distinctes.

- (1) Des poissons par centaines filaient et *nageaient* autour de l'étrave, emportés par le courant et le devançant à leur tour, allant de l'eau douce vers les eaux amères, et cette migration et ce départ étaient pleins de joie.
- (2) Puis, avec vigueur, il *nageait* la brasse vers le large... pas très loin.

Dans (1), il s'agit d'un prédicat monoargumental $X \sim$, alors que pour le sens (2), un second argument Y peut être sélectionné, ce qui nous amène à lui attribuer la forme propositionnelle $X \sim Y$.

La forme propositionnelle est ainsi un premier indice conduisant à distinguer deux acceptions du verbe **NAGER**, respectivement **Ia** et **Ib**, là où les dictionnaires ne traitent que d'un seul sens qu'il s'agisse d'un être humain ou d'un animal aquatique.

Nous verrons, dans la sous-section suivante, que l'explicitation de la forme propositionnelle entraîne de nombreuses conséquences pour le *definiens*.

2.2. STRUCTURE DU DEFINIENS

Les acceptions **Ia** et **Ib** ont reçu les paraphrases définitoires suivantes dans le RLF :

NAGERIa

L'animal $X \sim$:

- <CC label="se déplacer intentionnellement">L'animal X se déplace</CC>
- <PC role="milieu">en suspension dans l'eau</PC>
 - <PC role="gestuelle">par des mouvements de son corps</PC>

NAGERIb

L'individu $X \sim$ (en utilisant la technique Y) :

- <CC label="se déplacer intentionnellement">L'individu X se déplace</CC>
- <PC role="milieu">en suspension dans l'eau</PC>
 - <PC role="gestuelle">par des mouvements de son corps (conformes à la technique Y)</PC>
 - <PC role="background">ce mode de déplacement ayant été appris par X</PC>

Au prime abord, ces définitions peuvent paraître quelque peu déroutantes, notamment en raison des balises segmentant leurs contenus. Précisons que l'interface d'utilisateur prévue dans le RLF débarrassera le lecteur de ces marques. Il lira ainsi le texte de la définition de **NAGERIa** sous forme d'une phrase : « L'animal X se déplace en suspension dans l'eau par des mouvements de son corps ». Si nous avons choisi de garder les balises dans notre présentation, c'est parce que nous espérons d'une part, mettre en avant la structuration rigoureuse des paraphrases définitoires et de l'autre faciliter, le cas échéant, le renvoi à une composante définitionnelle précise. Autrement dit, elles nous permettront de naviguer plus aisément à travers les sous-réseaux

du lexique français, en suivant les liens qu'elles permettent d'établir entre les unités lexicales.

De fait, comme toutes les définitions du RLF, celles qui paraphrasent le sens de **NAGERIa** et de **NAGERIb** ont une structuration analytique : elles articulent une composante centrale (CC), qui exprime le genre prochain, et des composantes périphériques (PC) qui expriment la différence spécifique. La CC de chaque définition est porteuse d'une étiquette sémantique (*label*) qui correspond à « la forme normalisée du genre prochain de la définition qui établit l'appartenance de L à une classe de lexies d'au moins deux éléments » (Polguère 2011 : 202). Quant aux PC, le type d'information apportée par la composante est caractérisé par un attribut (*role*).

Notons que, dans le *definiens*, figurent obligatoirement toutes les variables actanciennes présentes dans la forme propositionnelle avec un typage sémantique, si nécessaire. Ainsi, les actants X de **NAGERIa** et de **NAGERIb** relèvent de deux types sémantiques différents, respectivement *animal* et *individu*. Il s'agit certes, d'une contrainte pour le lexicographe, mais également d'un outil précieux, puisque le rédacteur introduit ainsi dans la définition tous les participants de la situation, en spécifiant les rapports entre eux.

Toutes les PC, qui expriment la différence spécifique associée au genre prochain de la définition, doivent impérativement se rattacher directement à la CC. Il est interdit de hiérarchiser ces composantes entre elles, en leur adjoignant d'autres PC.

En plus de la CC et de PC, il est nécessaire, pour certaines lexies, d'introduire une composante présuppositionnelle. C'est ainsi que nous avons procédé en construisant la paraphrase de l'acception **NAGERVI** illustrée par l'exemple (3) :

(3) Débutante en latin, je *nage* complètement !!

L'unité lexicale **NAGERVI** a été définie de la façon suivante :

L'individu X ~

<PP role="contexte">Dans une situation complexe α que X doit maîtriser</PP>||
 <CC label="être dans un état psychologique">L'individu X ne comprend pas α </CC>
 <PC role="effet">ce qui fait que X éprouve une gêne</PC>

Selon notre analyse, l'information contextuelle concernant la complexité de la situation α que X est censé maîtriser, ne fait pas partie du sens verbal décrit. Il ne s'agit pas non plus d'un participant du fait dénoté par **NAGERVI**. C'est un contenu présupposé indispensable pour comprendre et utiliser cette lexie.

Le schéma ci-dessous (figure 2) réunit l'ensemble de composantes susceptibles d'apparaître dans la partie définitoire, en visualisant sa structuration. Il nous est impossible, dans le cadre de cet article, de citer l'inventaire complet des CP employées dans les définitions selon les attributs qu'elles portent. Celles dans la figure 2 ne sont que quelques exemples parmi d'autres, choisis dans le but d'illustrer notre propos sur la structure du *definiens*.

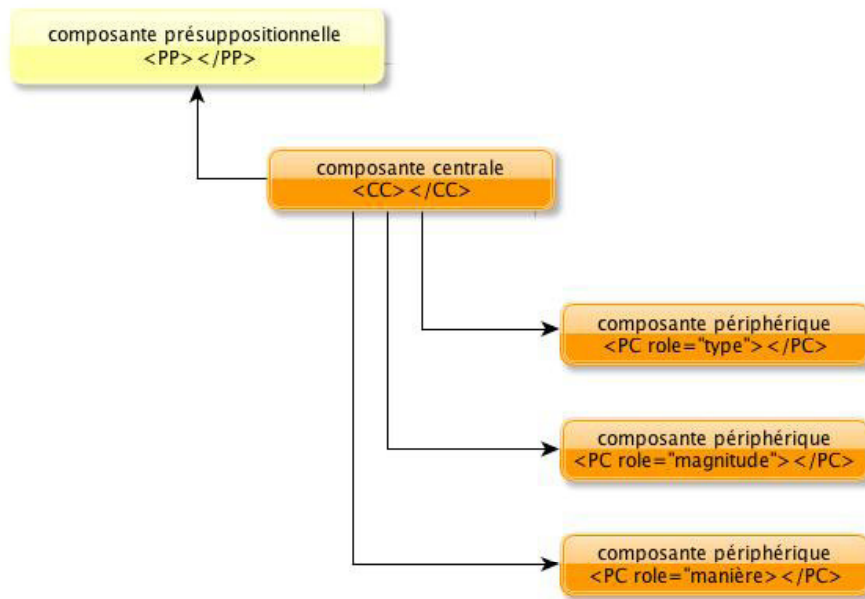


Fig. 2. Structure générale des définitions du RLF

3. DÉFINITIONS ET LIENS CONCEPTUELS

Dans cette section, nous expliquerons comment la structuration des définitions du RLF permet de reconstituer et de suivre les liens lexicaux de nature conceptuelle. Nous nous intéresserons tout particulièrement à ceux qui sont à la base des classes de sens (section 3.1) et à ceux qui relient les acceptions d'un vocable polysémique (section 3.2).

3.1. COMPOSANTES CENTRALES, COMPOSANTES PÉRIPHÉRIQUES ET CLASSES DE SENS

Commençons par le rôle que joue l'étiquette sémantique dans l'architecture d'une définition et bien au-delà de celle-ci. La CC du verbe **NAGER****la** porte la même étiquette « se déplacer intentionnellement », la même que celle que l'on retrouve dans la définition de **VOLER****la** présentée ci-dessous. L'exemple (4) illustre ce sens.

(4) Il regarda le ciel encore clair où *volaient* des martinets.

L'animal X ~ :

- <CC label="se déplacer intentionnellement">L'animal X se déplace</CC>
- <PC role="milieu">dans l'air</PC>
- <PC role="moyen">supporté par des parties spécifiques de son corps avec lesquelles il fait des mouvements qui lui permettent de se propulser</PC>

La classe de sens que constituent les verbes de déplacement français compte par ailleurs nettement plus de membres que le minimum requis par Polguère (*cf. supra*), puisqu'elle comprend les lexies telles que *courir*, *marcher*, *ramper* et bien d'autres.

Les principes définitionnels retenus dans le RLF permettent ainsi de situer la lexie décrite dans le sous-système conceptuel dont elle relève. Le paraphrasage rigoureux qu'impose une étiquette sémantique sert à connecter la lexie avec les autres membres de sa classe de sens. Nous avons pu constater, en effet, que les définitions de **NAGERIa** et de **VOLERIa** reposent sur une même paraphrase minimale. Dans une optique plus large, l'étiquette sémantique permet d'indiquer clairement la place qu'occupe une unité lexicale selon l'axe vertical de la hiérarchie des sens codés dans la langue. En remontant, on trouvera successivement des étiquettes de plus en plus vagues : « se déplacer », « être en mouvement », « exercer une activité », et ainsi de suite jusqu'à ce que le sommet de la hiérarchie soit atteint avec « fait » et « entité » qui se rejoignent sous l'étiquette commune « quelque chose ».

Parallèlement, les PC exprimant la différence spécifique d'une définition sont toujours caractérisées par un attribut. Une comparaison ne serait-ce que très rapide de **NAGERIa** et de **VOLERIa** permet de constater la présence des PC du même type dont le contenu véhicule les éléments de sens propre à chacune des deux lexies. Les attributs (milieu, gestuelle dans le cas de **NAGERIa**) qui caractérisent les PC apparaissent, bien que dans des configurations différentes, dans les définitions des verbes de la même classe de sens. Ainsi, pour définir la lexie **VOLERIa**, on recourt à la PC milieu, alors que la "gestuelle" est remplacée par "moyen" pour inclure le vol plané des oiseaux.

3.2. STRUCTURE DES DÉFINITIONS ET TRAITEMENT DE LA POLYSÉMIE DANS LE RLF

Il n'est pas rare que les lexies d'un même vocable polysémique appartiennent à des classes de sens bien différentes. Il suffit de penser au verbe **ROULER** dont la première acception **ROULERI1** dénote un déplacement non intentionnel (exemple 5), **ROULERI2** de (6) est un causatif de déplacement, alors qu'avec **ROULERII** dans (7), il s'agit plutôt d'un mouvement corporel.

(5) Les pierres *roulent* sous les pieds ; on se cramponne ; on se rattrape aux branches ; mon sabre devient un alpenstock.

(6) Un paysan *roule* un tonneau.

(7) Il grince des dents et *roule* des yeux féroces d'assassin de film muet.

Avant de présenter le traitement lexicographique de la polysémie dans le réseau lexical décrit par le RLF, quelques précisions terminologiques s'imposent.

Nous désignerons désormais les unités lexicales (lexies, acceptions, sens) d'un vocable polysémique sous le nom de **co-polysèmes**.

Les liens reposant sur des composantes sémantiques communes aux co-polysèmes d'un vocable polysémique seront appelés **ponts sémantiques**⁶.

L'un des co-polysèmes constitue l'unité lexicale de base (ULB), c'est-à-dire le sens auquel tous les autres sont connectés par des ponts sémantiques directement ou indirectement, *via* une autre acception.

La position résolument synchronique du RLF est décisive pour le traitement de la polysémie. Il faut en effet que les ponts sémantiques entre les co-polysèmes soient accessibles à l'utilisateur sans détour étymologique. Prenons le cas du verbe **ROULER** dans la phrase (8) ci-dessous :

(8) Nous étions en plein procès avec l'ancien propriétaire qui nous *avait roulés* et nous allions perdre.

Dans les dictionnaires⁷ qui décrivent la polysémie en s'appuyant sur le « parcours » étymologique du vocable, la lexie **ROULER** de l'exemple (8) est un co-polysème de **ROULER** employé dans (5), (6) et (7). Dans le RFL cependant, on sera amené à les traiter en termes d'homonymie, puisqu'à moins d'une construction conceptuelle très complexe et artificielle, il est impossible de construire un pont sémantique en synchronie susceptible de relier le sens de l'exemple (8) aux précédents. Il convient par conséquent d'en faire deux articles séparés⁸, respectivement **ROULER1** et **ROULER2**.

Les composantes définitionnelles sont l'outil essentiel pour relier les co-polysèmes. La première étape consiste à identifier l'ULB, car c'est elle qui constitue le point de départ du raisonnement structurant toutes les autres acceptions. Autrement dit, c'est en partant de l'ULB et de sa paraphrase définitoire que le lexicographe – et à sa suite l'utilisateur du RLF – pourra construire un premier pont sémantique, puis un deuxième et tous ceux qui le mèneront d'une acception à l'autre pour comprendre la rationalité qui sous-tend la structuration d'un vocable polysémique. Si la définition de l'ULB est le point de départ, il faut que le pont sémantique soit solidement ancré à l'autre bout, c'est-à-dire du côté du co-polysème cible. La seconde tâche dans le traitement de la polysémie est donc d'élaborer des paraphrases définitoires qui explicitent les liens entre eux.

Dans la section 3.2, nous avons présenté les lexies **NAGER1a** et **NAGER1b** reliées par la même CC et par la PC 'milieu'. Cependant, les connexions sont parfois plus complexes. Observons les définitions de deux acceptions qui appartiennent à la très

⁶ Nous adoptons ainsi les notions proposées et définies par Mel'čuk, Clas et Polguère (1995 : 157).

⁷ Cf. à titre d'exemple *Le Petit Robert* (2011) ou bien le site du *Centre National des Ressources Textuelles et Lexicales* (<http://www.cnrtl.fr/definition/roule>).

⁸ L'enquête que nous avons menée auprès de francophones natifs indique clairement qu'il n'est pas possible de rattacher **ROULER2** à **ROULER1** *via* la locution *rouler quelqu'un dans la farine*. Il semblerait en effet qu'un locuteur contemporain puisse utiliser le verbe *rouler* avec le sens de *tromper* tout en ignorant la locution en question.

riche polysémie du vocable **ROULER**. En effet, **ROULER^{I1}** et **ROULER^{I2}**, ont été illustrées, au début de la présente section, respectivement par les exemples (5) et (6).

ROULER^{I1}

$X \sim \text{sur } Y$

- <CC label="être en déplacement">X est en déplacement</CC>
- <PC role="causateur">par des mouvements autour de son axe</PC>
- <PC role="milieu">sur le support Y</PC>

ROULER^{I2}

$W \sim X \text{ sur } Y$

- <CC label="déplacer">W déplace X</CC>
- <PC role="manière">en faisant rouler^{I1} X</PC>

La lexie **ROULER^{I1}**, ULB du vocable **ROULER**(1), est sémantiquement la plus simple parmi les co-polysèmes et elle sert de base pour la construction de ponts sémantiques. Le premier d'entre eux, celui qui la relie avec le sens **I2**, est une inclusion. **ROULER^{I2}** est plus riche, puisqu'il lexicalise une relation causale, en sélectionnant un actant agentif. Autrement dit, il inclut le sens de base, en y ajoutant du contenu informationnel.

Il n'est pas rare qu'au fur et à mesure que l'on s'éloigne de l'ULB, les ponts deviennent plus difficiles à identifier. Il est parfois impossible de détecter des composantes communes avec la lexie de base. **ROULER^{II}** de l'exemple (7) ci-dessous semble effectivement bien différent de celui que nous venons de définir pour la phrase (5).

Le lien qui relie **ROULER^{II}** à l'ULB est de toute autre nature que dans le cas du causatif **I2**. L'hypothèse que nous avons adoptée consiste à admettre que le pont entre **ROULER^{I1}** et **II** employée dans l'exemple (7) se construit par métaphore. Contrairement à l'ensemble de dictionnaires que nous avons pu consulter, le RLF ne se contente pas de laisser à l'utilisateur le soin d'identifier les liens métaphoriques. Une métaphore lexicalisée étant partie intégrante du sens, elle est systématiquement explicitée dans la définition-cible à travers une PC spécifique porteuse de l'attribut 'métaphore' :

ROULER^{II}

L'individu X ~ la partie de son corps X1

- <CC label="remuer">L'individu X remue la partie X1 de son corps</CC>
- <PC role="manière">par des mouvements circulaires</PC>
- <PC role="métaphore">comme si X1 *roulait*.I</PC>

Construire des ponts sémantiques en définissant est ainsi une tâche qui génère une double contrainte. D'une part, en paraphrasant le sens de départ sans perdre de vue les co-polysèmes auxquels il doit nous mener par la suite. De l'autre, il faut éviter la redondance dans les définitions des sens ciblés. Cette discipline vaut pourtant l'effort qu'elle requiert dans la mesure où les mécanismes des évolutions sémantiques à l'ori-

gine de la polysémie produisent, certes, des effets spécifiques à chaque langue, mais sont identiques quant à leur nature.

4. DÉFINITIONS ET LIENS PARADIGMATIQUES ET SYNTAGMATIQUES

Jusqu'à présent, nous avons employé le terme de paraphrase, en la qualifiant de définitoire, comme synonyme de définition. Nous adopterons à présent une perspective plus large pour nous occuper des liens lexicaux qui autorisent le paraphrasage tel qu'il est placé au cœur du modèle Théorie Sens-Texte. Par paraphrasage, on entend la synonymie entre les phrases (*textes*) qui expriment le même sens, par exemple entre (9) et (10) :

- (9) Jean *nage très bien*.
 (10) Jean est un *très bon nageur*.

Nous présentons le traitement des relations paradigmatiques dans le RLF dans la sous-section (4.1), pour passer ensuite, en (4.2), aux liens syntagmatiques.

4.1. DÉFINITIONS ET LIENS PARADIGMATIQUES

Dans la section 3, nous avons présenté les définitions des lexies **NAGERIa** et **NAGERIb** qui dénotent respectivement le déplacement d'un animal aquatique, typiquement un poisson, et celui qui peut être maîtrisé par un être humain. Ce choix peut paraître étonnant dans la mesure où les dictionnaires aussi bien traditionnels qu'informatisés n'y voient habituellement qu'un seul sens.

Certes, les deux verbes ont des structures actanciennes différentes, traduites dans deux formes propositionnelles distinctes dans le *definiendum*. Cela n'est pourtant pas une raison suffisante d'en faire deux lexies-vedettes. On pourrait en effet considérer que le second actant Y de **NAGERIb** (style de natation) correspond à une composante sémantique faible qui n'est pas systématiquement actualisée dans le discours, notamment lorsque X est un animal⁹. Dans cette optique, on aurait affaire à une seule acception **NAGERI** dont la forme propositionnelle serait la suivante :

L'être animé X ~ (en utilisant la technique Y).

En adoptant cette hypothèse, il faudrait proposer une paraphrase définitoire suffisamment large pour couvrir les deux cas. Cette solution a été envisagée à une étape antérieure de notre travail avec la définition suivante :

<CC label="se déplacer intentionnellement">L'être animé X se déplace</CC>
 <PC role="milieu">en suspension dans l'eau</PC>
 <PC role="gestuelle">par des mouvements de son corps (conformes à la technique Y)</PC>

⁹ Remarquons que le second actant Y correspond à une composante faible même lorsque l'on sépare les acceptions **NAGERIa** et **NAGERIb**, dans la mesure où un être humain peut très bien se déplacer en suspension dans l'eau sans que ses gestes forment un style quelconque.

Rappelons cependant que la démarche définitoire adoptée dans le cadre du RLF vise un triple objectif. Il s'agit de :

- i. décrire le sens,
- ii. identifier le sous-réseau des unités lexicales auxquelles la lexie-vedette est connectée,
- iii. situer la lexie-vedette dans le graphe que forme le lexique du français.

Si une acception commune **NAGERI** est satisfaisante en ce qui concerne le postulat i., elle est loin de répondre aux exigences ii. et iii. En effet, elle ne prend pas en compte les liens paradigmatiques et syntagmatiques (sous-section 4.2) qui structurent le lexique de la langue.

Or, en ce qui concerne les liens lexicaux, le verbe **NAGERIa** de l'exemple (1) et **NAGERIb** de (2) forment deux réseaux lexicaux sensiblement différents. Il suffit de penser aux exemples (9) et (10) dans la partie introductive de cette section : **NAGEUR** et **NAGEUSE** sont bien des noms de l'actant X mentionné dans la définition de **NAGERIb**, mais pas de celui de **NAGERIa**. Ainsi, dire d'un poisson rouge qu'il est un excellent nageur est toujours possible, mais relève plutôt d'une plaisanterie.

Le tableau 1 réunit les lexies connectées avec les acceptions **NAGERIa** et **NAGERIb** par des relations paradigmatiques. On désignera ces lexies sous le nom de dérivés sémantiques.

Tableau 1. Dérivés sémantiques des lexies **NAGERIa** et **NAGERIb**

Type de lien paradigmatique	NAGERIa	NAGERIb
quasi-synonymes	—	CRAWLER FLOTTER, BARBOTER, PATAUGER
nominalisation	NAGEIa	NAGEIb
adjectivation	NATATOIREa	NATATOIREb
noms pour l'actant X	<i>typiquement</i> : POISSON	NAGEURA, NAGEUSEa
nom pour l'actant Y	—	NAGEII ; BRASSE, BRASSE PAPILLON, CRAWL, DOS BRASSÉ, DOS CRAWLÉ, NAGE INDIENNE, PAPILLON

Il est frappant de constater la disparité quantitative des dérivés sémantiques de ces deux acceptions. Le sens « humain » développe dans la langue un réseau de connexions nettement plus vaste que le sens « animal ». Ce dernier donne lieu à une conversion nominale, adjectivale et à un nom d'actant X.

Les lexies nominales **NAGEIa** et **NATATOIREa** sémantiquement construites sur **NAGERIa** dénotent le même fait que leur source de dérivation. Les manipulations que nous effectuons sur l'exemple (11) conduisent aux phrases (12) et (13) qui expriment un sens similaire.

- (11) Les poissons-pilotes ne guident pas les requins, comme on le pensait autrefois, ils profitent du flux hydrique créé par la *nage* du requin. (Internet)
- (12) Les poissons-pilotes ne guident pas les requins, comme on le pensait autrefois, ils profitent du flux hydrique créé par le requin qui *nage*.
- (13) Les poissons-pilotes ne guident pas les requins, comme on le pensait autrefois, ils profitent du flux hydrique créé par l'activité *natatoire* du requin.

Pour éviter la redondance, tout en indiquant clairement le lien paradigmatique, nous proposerons dès lors les définitions suivantes de ces dérivés sémantiques :

NAGEIa

~ de l'animal X

<CC label="action de se déplacer intentionnellement">Fait que l'animal X **NAGEIa**</CC>

NATATOIREa

[X] ~

<CC label="relatif au fait de se déplacer intentionnellement">Relatif au fait que l'animal X **NAGEIa**</CC>

De la même manière, la lexie **NAGERIb** servira de base pour définir ses dérivés sémantiques réunis dans la colonne de droite de notre tableau. Faute de place, nous laisserons au lecteur le plaisir de cet exercice.

Cette démarche rigoureuse conduit par ailleurs à éviter des rapprochements trop intuitifs. Ainsi, le nom **NATATION** ne figure pas parmi les dérivés sémantiques de **NAGERIb**, car on ne saurait définir ces lexies en partant de ce verbe. Pour s'en convaincre, il suffit de rappeler l'exemple (2) : **NAGERIb** la brasse vers le large, même vigoureusement, ne revient point à faire de la natation. La lexie **NATATION** semble connectée par un lien paradigmatique à **NAGERII** qui est un verbe de réalisation que nous qualifierions de sportive, elle-même sémantiquement construite sur **NAGERIb**.

Le travail de construction de définitions poursuivi dans le RLF est donc en fait un incessant va-et-vient entre la lexie-vedette et ses dérivés sémantiques. D'une part, la définition ouvre l'accès aux lexies connexes dans la mesure où elle permet à l'utilisateur de prévoir les dérivés sémantiques. On comprend dès lors à quel point il est important de spécifier, dans la paraphrase définitoire, tous les actants représentés dans la forme propositionnelle (voir section 3) et les rapports entre eux. C'est ainsi qu'il est possible d'identifier le sous-réseau d'unités lexicales dont le locuteur a besoin pour désigner un fait donné par la phrase qu'il pourra choisir parmi celles qui conviennent.

D'autre part, le lexicographe parcourt régulièrement le chemin en sens inverse, en allant du dérivé vers la lexie-vedette. En effet, une définition des lexies **NAGEa** et **NAGEb** ne peut se faire qu'en fonction de celles qui constituent leurs sources de dérivation sémantique. Cela signifie *a contrario* qu'en construisant la définition de la lexie source (en l'occurrence celles de **NAGERIa** et de **NAGERIb**), le lexicographe ne perd pas de vue le fait qu'elle sera au cœur des paraphrases définitoires de ses dérivés sémantiques.

4.2. DÉFINITIONS ET LIENS SYNTAGMATIQUES

Une définition du RLF doit également permettre à l'utilisateur d'explorer les liens syntagmatiques propres à la lexie décrite. Or, l'analyse de la combinatoire syntaxique de **NAGERIa** et de **NAGERIb** fournit des arguments solides en faveur d'un traitement séparé des deux sens, *ipse eo* de deux paraphrases définitoires distinctes.

On constate tout d'abord que le réseau syntagmatique de **NAGERIb** est nettement plus riche que celui de l'ULB **NAGERIa**. En ce qui concerne Jean, excellent nageur des exemples (9) et (10), il est possible d'apprécier ses talents en disant qu'il nage comme un poisson. Certes, on pourra toujours dire d'un serpent de mer qu'il **NAGEIa** comme un poisson. Il s'agira néanmoins non pas d'une évaluation positive mais d'une comparaison de modes de déplacement.

Cela signifie que l'unité lexicale **NAGERIb** constitue la base des collocations formées avec des collocatifs *comme un poisson, comme un boulet, comme un caillou, comme un fer à repasser*, etc., incompatibles avec le sens **NAGERIa**. Or, une définition du RLF doit compter avec ce type de liens. En ce qui concerne **NAGERIb**, les CP "gestuelle" et "background" indiquent que le degré de maîtrise des gestes lié tant à l'apprentissage qu'à la pratique trouve une expression dans la langue. De même, le caractère inné de ce mode de déplacement chez un poisson bloque la possibilité de former la collocation *apprendre à nagerIa*, alors qu'elle n'est pas rare dans le cas de **NAGERIb**.

4.3. DÉFINITIONS, LIENS LEXICAUX ET STRUCTURE DU SOUS-RÉSEAU LEXICAL

La troisième exigence à laquelle répondent les définitions du RLF est celle de identifier le sous-réseau de la lexie-vedette. Avant de conclure nos propos, il est temps de présenter les graphes des deux lexies qui nous ont particulièrement intéressées dans ces pages. Les arcs continus marquent les liens paradigmatiques, ceux en pointillés indiquent les relations syntagmatiques. La disparité entre la densité des liens paradigmatiques et syntagmatiques dans les graphes (3) et (4) est désormais frappante.

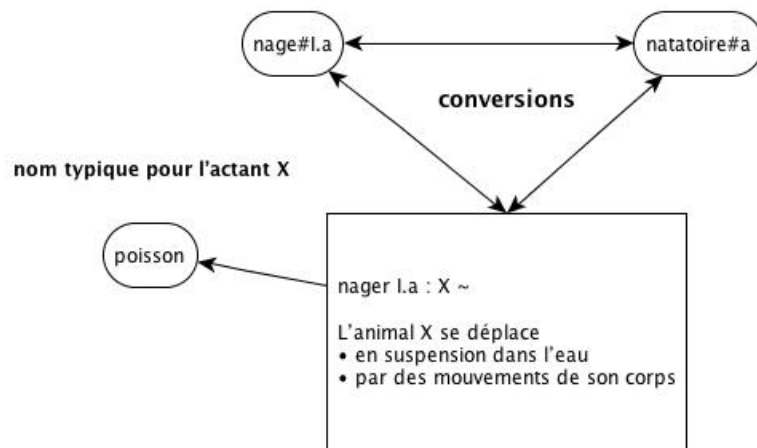


Fig. 3. Réseau de relations lexicales de NAGERIa

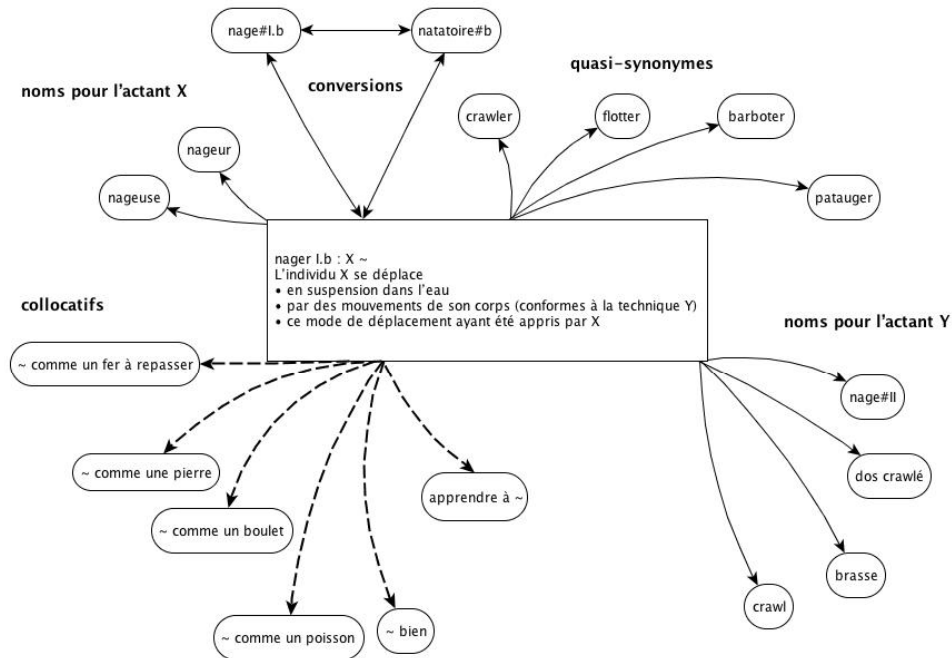


Fig. 4. Réseau de relations lexicales de NAGERIb

5. POUR CONCLURE

Définir en suivant fidèlement les liens lexicaux quelle que soit leur nature (cf. sections 3 et 4) est un défi qui – à notre connaissance – n’a pas encore été relevé dans une ressource offrant une large couverture du français. Au fil de ces pages, nous avons vu le nombre de contraintes à respecter pour construire les définitions qui tissent un réseau de liens reproduisant la structure du lexique, en l’occurrence celui du français. Elles sont, certes, très fortes mais débouchent en réalité sur une richesse d’informations inestimables pour tout utilisateur. Le linguiste y trouvera, entre autres, de quoi explorer les phénomènes de polysémie. Un élève de langue maternelle pourra explorer le réseau à la recherche du « mot juste » qui lui permettra d’éviter le verbe *être* omniprésent, mais invariablement mal vu par son professeur de français. Les apprenants étrangers de différents niveaux le parcourront à la recherche du collocatif manquant. À chacun son réseau ... lexical.

BIBLIOGRAPHIE

- LUX-POGODALLA, Veronika ; POLGUÈRE, Alain (2011): *Construction of a French Lexical Network : Methodological Issues. Proceedings of the First International Workshop on Lexical Resources, WoLeR 2011. An ESSLLI 2011 Workshop*. Ljubljana, 54-61.
- MCCARTHY, Michael (1999): « What constitutes a basic vocabulary for spoken communication ? ». *English and Literature* 1: 233-249.
- MEL’ČUK, Igor Aleksandrovič et alii (1984, 1988, 1992, 1999): *Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques*. Volumes I–IV. Montréal : Les Presses de l’Université de Montréal.
- MEL’ČUK, Igor Aleksandrovič (1996): « Lexical Functions: A Tool for the Description of Lexical Relations in the Lexicon ». In: WANNER, Leo [ed.]. *Lexical Functions in Lexicography and Natural Language Processing*. Amsterdam/Philadelphie : John Benjamins, 37-102.
- MEL’ČUK, Igor Aleksandrovič (1997): *Vers une linguistique Sens-Texte*. Leçon inaugurale. Paris : Collège de France.
- MEL’ČUK, Igor Aleksandrovič ; CLAS, André ; POLGUÈRE, Alain (1995): *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve : Duculot.
- MEL’ČUK, Igor Aleksandrovič ; CLAS, André ; POLGUÈRE, Alain (2007): *Lexique actif du français. L’apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*. Bruxelles : De Boeck.
- MILIČEVIĆ, Jasmina (2006): « A Short Guide to the Meaning-Text Linguistic Theory ». *Journal of Koralex* 8: 187-233.
- POLGUÈRE, Alain (2011a): « Classification sémantique des lexies fondée sur le paraphrasage ». *Cahiers de Lexicologie* 98: 197-211.
- POLGUÈRE, Alain (2011b) : « Perspective épistémologique sur l’approche linguistique Sens-Texte ». In: *L’architecture des théories linguistiques, les modèles et leurs interfaces*. Mémoires de la Société de Linguistique de Paris, t. XX: 79-114.
- POLGUÈRE, Alain (à paraître): « Lexicographie des dictionnaires virtuels ». In: *Meaning, Text and Other Exiting Things. Hommages à Igor Mel’čuk*.

- POLGUÈRE, Alain ; SIKORA, Dorota (à paraître): « Modèle lexicographique de croissance de vocabulaire fondé sur un processus aléatoire, mais systématique ». In : *Enseigner le lexique. Pratiques sociales objets à enseigner et pratiques d'enseignement*. Metz : AiRDF.
- SELVA, Thierry ; VERLINDE, Serge ; BINON, Jean (2003): « Vers une deuxième génération de dictionnaires électroniques ». *Traitement Automatique des Langues (T.A.L.)* 44(2): 177-197.
- WIERZBICKA, Anna (1996): *Semantics. Primes and Universals*. Oxford – New York : Oxford University Press.

Ressources électroniques

- URL: <<http://www.cnrtl.fr>>. (Last accessed at: 8 juin 2012)
- URL: <<http://www.frantext.fr>>. (Last accessed at: 8 juin 2012)
- URL: <<http://olst.ling.umontreal.ca/dicopop/>>. (Last accessed at: 8 juin 2012)
- URL: <<http://olst.ling.umontreal.ca/dicouebe/>>. (Last accessed at: 8 juin 2012)

