

# Notions de sémantique lexicale et leur application dans des ressources lexicales et terminologiques

Marie-Claude L'Homme
Observatoire de linguistique Sens-Texte (OLST)
Université de Montréal

mc.lhomme@umontreal.ca

www.olst.umontreal.ca





## Programme ...

#### 1. Jour 1

- WordNet
- FrameNet
- « Unité lexicale »
- Polysémie
- (si le temps) Dicouèbe

#### 2. Jour 2

- DiCoInfo et DicoEnviro
- « Terme »
- Structure argumentale
- Quelques liens lexicaux



## Programme Jour 1

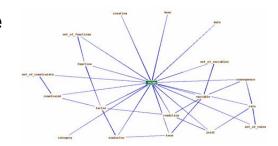
- 1. Quelques ressources lexicales
  - Ressources lexicales : WordNet et FrameNet, Dicouèbe
    - Concepts théoriques régissant ces ressources
      - Unité lexicale
      - Polysémie



## 1. Quelques ressources lexicales et terminologiques

#### WordNet

- Une ressource lexicale pour l'anglais
  - Des versions dans <u>de très nombreuses langues</u>
- S'appuie sur des hypothèses psycholinguistiques selon lesquelles le lexique est organisé en réseau (Miller et al. 1999)
- Principes structurants (ex. green)
  - Unités lexicales organisées en synsets
  - Noms : hyperonymie (taxinomie), méronymie
  - Verbes : hyperonymie
  - Adjectifs et adverbes : antonymie
  - Tous : liens entre formes





## 1. Quelques ressources lexicales et terminologiques

#### FrameNet

- Une ressource lexicale pour l'anglais
  - Des versions dans <u>d'autres langues</u>
- S'appuie sur l'hypothèse selon laquelle le lexique se construit sur des connaissances d'arrière-plan (background knowledge) : Frame Semantics (Fillmore 1975; Fillmore et al. 2003)
- Principes structurants et éléments descriptifs :
  - Cadre sémantique : scénario conceptuel évoqué par des unités lexicales (ex. enseigner; teach)
  - Éléments d'un cadre (frame elements)
  - Unité lexicale
  - Annotations
  - Relations entre cadres

Frame



- « Unité lexicale » : unité linguistique résultant de l'association d'une forme lexicale ayant une fonction syntaxique et d'un sens lexical ≠ mot
  - Problème de définition de la notion de « mot »
  - Problème de forme (délimitation graphique infructueuse)
    - La pomme de terre est sur la table.
  - Problème de sens
    - Je connecte la souris à l'ordinateur.
    - Le chat poursuit la souris.
  - Problème de sens et de forme
    - Pré- dans préavis n'est pas un mot



- « Unité lexicale »
  - Cruse (1986)
  - Délimitation syntagmatique
  - une unité lexicale doit correspondre à au moins une composante sémantique
  - (i) Composante sémantique : toute composante d'une phrase qui véhicule un sens qui, en se combinant avec le sens des autres unités de la phrase, participe au sens global de cette phrase.
    - Orange dans l'orange est sur la table correspond à une composante sémantique
    - Pomme dans la pomme de terre est sur la table ne correspond pas à une composante sémantique



#### « Unité lexicale »

- Délimitation syntagmatique
- Test 1 : La participation d'une composante sémantique au sens d'une phrase est la même dans d'autres phrases.

L'orange est sur la table.

Je mange une orange chaque jour.

Le col de cette chemise est bleu.

Le ciel est bleu aujourd'hui.

\*Les cols bleus de Montréal décident encore une fois de prendre des mesures.



#### « Unité lexicale »

- Délimitation syntagmatique
- Test 2 : Le remplacement d'une UL par une autre dans d'autres phrases distinctes crée un contraste sémantique récurrent.

L'orange est sur la table. La pomme est sur la table.

La saison des oranges commence. La saison des pommes commence.

Je mange une orange chaque jour. Je manque une pomme chaque jour.

Je mange une pomme de terre chaque jour. La pomme de terre est sur la table.

Je mange une carotte chaque jour. La carotte est sur la table.



#### « Unité lexicale »

- Délimitation syntagmatique
- Test 2 : Le remplacement d'une UL par une autre dans d'autres phrases distinctes crée un contraste sémantique récurrent.

Les cols bleus de Montréal décident encore une fois de prendre des mesures.

Les cols blancs de Montréal décident encore une fois de prendre des mesures.

\*Les cols rouges de Montréal décident encore une fois de prendre des mesures.

Ma chemise a un col bleu.

Ma chemise a un col blanc.

Ma chemise a un col rouge.



- « Unité lexicale »
  - Cruse (1985)
  - Délimitation syntagmatique
  - (ii) une unité lexicale doit correspondre à au moins un mot
  - Mot : plus petit élément de la phrase qui peut changer de place dans celle-ci dans affecter sa grammaticalité (le sens, toutefois, est généralement affecté).
    - Compilation dans Je procède à la compilation du programme est un mot.
    - Pré- dans je procède à la précompilation du programme n'est pas un mot.



- « Unité lexicale »
  - Cruse (1986)
  - Délimitation syntagmatique
  - une unité lexicale doit correspondre à au moins un mot Jean voit Pierre.

Pierre voit Jean.

Pierre, Jean voit.

Jean revoit Pierre.

Pierre voit re- Pierre.

L'unité la plus grande qui ne peut accepter d'insertion.

Je procède à la précompilation du programme.

Je procède (dès maintenant) à la précompilation (rapide et efficace) du (nouveau) programme.



- Unités lexicales dans les ressources
  - Les ressources lexicales tranchent souvent avec une pratique lexicographique classique en ce sens qu'elles admettent comme unités lexicales des séquences différentes
    - Unité lexicale simples : lexème
    - Unité lexicale complexe :
      - phrasème complet (expression idiomatique) : coup de foudre, levée de boucliers
      - □ quasi-phrasème : *centre commercial, boîte aux lettres*
  - Exemples (<u>Dicouèbe</u>):
    - Jeter un coup d'œil à la liste des lexies; levée de boucliers
  - WordNet : kick the bucket



- Unités lexicales dans les ressources
  - Les ressources lexicales tranchent souvent avec une pratique lexicographique classique en ce sens que les unités lexicales décrites sont celles associées à des classes ouvertes : noms, verbes, adjectifs, adverbes, préposition (de sens plein et non les emplois imposés par des contraintes grammaticales).
  - FrameNet :
    - jeter un coup à la nomenclature (lettre a); About (topic)



## Programme Jour 1

- 1. Quelques ressources lexicales
  - Ressources lexicales : WordNet et FrameNet, Dicouèbe
    - Concepts théoriques régissant ces ressources
      - Unité lexicale
      - Polysémie



- Polysémie
  - Exercice <u>vert</u>; <u>clou</u>
- Premier exercice de distinction de sens
  - Vert



#### Polysémie (tests)

■ Synonymie: pour une forme lexicale X, s'il existe un synonyme pour une de ses occurrences, mais que ce synonyme n'est pas valable pour une seconde occurrence de cette forme dans un autre contexte; alors, nous avons deux sens.

C'est donc en toute logique que la France a décidé d'accélérer le pas dans la voie des carburants **vert**s.

Des carburants écologiques.

Des carburants environnementaux.

Cette personne a les yeux verts.

\*les yeux écologiques

\*les yeux environnementaux



#### Polysémie

■ Antonymie: pour une forme lexicale X, s'il existe un antonyme pour une de ses occurrences, mais que cet antonyme n'est pas valable pour une seconde occurrence de cette forme dans un autre contexte; alors, nous avons deux sens.

C'est donc en toute logique que la France a décidé d'accélérer le pas dans la voie des carburants **vert**s.

Des carburants polluants.

Des carburants fossiles.

Je pense que Paris-Rive gauche sera très avancé, avec les choix architecturaux pour l'université, la création de logements supplémentaires, nouveaux espaces **verts**.

\*Des espaces polluants. \*Des espaces fossiles.



#### Polysémie

■ Dérivation différentielle (relation paronymique chez Cruse 1986): Pour une forme lexicale X, s'il existe un dérivé rattaché à une de ses occurrences, mais que ce lien n'est pas valable pour une seconde occurrence dans un autre contexte; alors, nous avons deux sens.

Toute sa vie, Jeanne a élevé des enfants.

\*a fait l'élevage d'enfants.

\*éleveuse

Toute sa vie, Jeanne a élevé des cochons.

a fait l'élevage des cochons éleveuse de cochons



#### Polysémie

#### Dérivation différentielle

Des boulons, des **clous** et des morceaux de métal, ajoutés à la bombe, ont été retrouvés sur place.

... parce que les objets déposés en gage étaient suspendus à des clous.

Clouer, clouage

Le cuir est sculpté, découpé au laser, rivé de **clous**, comme le sac Domino de Sonia Rykiel

\*clouer



#### Polysémie

- Autres tests
- Cooccurrence compatible (Mel'cuk et al. 1995): Pour une forme lexicale X, si on peut combiner des cooccurrents dans une phrase normale, alors nous avons un sens.
- Si, pour la lexie potentielle L (....), on peut construire une phrase normale à cooccurrence compatible; alors L ne doit pas être scindée.

Des espaces et des quartiers verts.

\*Des espaces et des yeux verts.



#### Polysémie

- Autres tests
- Cooccurrence différentielle (Mel'cuk et al. 1995): Pour une forme lexicale X, si la combinaison des cooccurrents donne lieu à une phrase anormale, alors nous avons deux sens.
- Si pour la lexie potentielle L (....), on peut dégager deux ensembles disjoints de cooccurrents, alors L doit être scindée.

?Des carburants et des espaces verts.

Des formes d'énergie ou technologiques vertes.



#### Polysémie

- Autres tests
- Autres liens paradigmatiques: pour une forme lexicale X, si on peut identifier son appartenance à une série paradigmatique pour une de ses occurrences, mais que cette appartenance n'est pas valable pour une seconde occurrence de cette forme dans un autre contexte; alors, nous avons deux sens.



#### Polysémie

- Autres tests
- Autres liens paradigmatiques

Son logo est **vert**, ceux de France 2 et France 3 restent respectivement rouge et bleu

Vert, rouge, bleu \*écologique,

Le parti conservateur, ainsi que les deux partis d'opposition représentés au Parlement - les **verts** et les sociaux-démocrates - ont dénoncé une manoeuvre risquant de bloquer tout dialogue avec Prague.

Vert, social-démocrate, conservateur \*jaune, marron



## Quelques ressources lexicales

#### Dicouèbe

- Une ressource lexicale pour le français
- S'appuie sur l'hypothèse selon laquelle le lexique est structuré : Lexicologie explicative et combinatoire (Mel'čuk et al. 1995)
- Principes structurants et éléments descriptifs (une partie):
  - Unité lexicale (ex. avocat)
  - Structure actancielle (ex. avocat)
  - Lien lexicaux
  - Régime





- Polysémie dans les ressources lexicales
  - Comment rend-on compte de la distinction des sens dans :
    - WordNet (mouse; president; green)
    - FrameNet (green, impact, warm)
    - Dicouèbe (abandon, bijou, canard)



## Programme Jour 2

- 1. Des ressources terminologiques
  - DiCoInfo et DicoEnviro
- 2. Notion de « terme »
- 3. Structure argumentale
- 4. Quelques liens lexicaux



### DiCoInfo et DiCoEnviro

DicoInfo, Dictionnaire fondamental de l'informatique et de l'Internet

Développement commencé vers 2002

Entrées en anglais, français et espagnol

Développement en français plus abouti

DiCoEnviro, Dictionnaire fondamental de l'environnement

Développement commencé vers 2010

Entrées en anglais, français, espagnol et portugais





## DiCoInfo et DiCoEnviro : objectifs

- Fournir une description détaillée du fonctionnement linguistique des termes
  - Structure argumentale, annotations contextuelles, liens lexicaux
- Donner une description détaillée qui reflète le fonctionnement des termes dans les corpus
  - Propriétés combinatoires
  - Préférences lexicales
- Tenir compte de distinctions sémantiques fines
- Fournir de longues listes de liens lexicaux



## DiCoInfo et DiCoEnviro : exemples

- courriel
- terre



#### Définition de « terme »

L'effet de serre est un processus naturel résultant de l'influence de l'atmosphère sur les différents flux thermiques contribuant aux températures au sol d'une planète. La prise en compte de ce mécanisme est nécessaire pour expliquer les températures observées à la surface de la Terre et de Vénus. Dans le système solaire l'essentiel de l'énergie thermique reçue par une planète provient du rayonnement solaire et, en l'absence d'atmosphère, une planète rayonne idéalement comme un corps noir, l'atmosphère d'une planète absorbe et réfléchit une partie de ces rayonnements modifiant ainsi l'équilibre thermique. Ainsi l'atmosphère isole la Terre du vide spatial comme une serre isole les plantes de l'air extérieur. L'expression effet de serre résulte d'une analogie entre l'atmosphère et les parois d'une serre. Son usage s'est étendu dans le cadre de la vulgarisation du réchauffement climatique causé par les gaz à effet de serre qui bloquent et réfléchissent une partie du rayonnement thermique. Or le bilan thermique d'une serre s'explique essentiellement par une analyse de la convection et non du rayonnement : la chaleur s'accumule à l'intérieur de la serre car les parois bloquent les échanges convectifs entre l'intérieur et l'extérieur. Aussi, le terme scientifique, utilisé par la communauté des climatologues pour décrire l'influence des gaz à effet de serre, composants de l'atmosphère bloquant le rayonnement infrarouge, sur le bilan thermique de la Terre, est forçage radiatif. Les températures terrestres résultent d'interactions complexes entre les apports solaires perturbés par les cycles de l'orbite terrestre, de l'effet albédo de l'atmosphère, des courants de convection dans l'atmosphère et les océans, du cycle de l'eau et le forçage radiatif de l'atmosphère notamment.



#### Définition de « terme »

#### Le terme, un type d'unité lexicale

- Entités liées au domaine (ex. environnement)
  - Objets concrets (ex. éolienne, planète)
  - Représentation (ex. scénario)
  - Substances (gaz, méthane, dioxyde de carbone)
- Unités prédicatives (nature des arguments)
  - réfléchir (planète réfléchit rayonnement)
  - réchauffement du climat
- Familles morphologiques
  - Polluer, pollution, polluant, pollueur
- Liens paradigmatiques
  - Réchauffer, changer, refroidir



- Unités prédicatives et actants sémantiques
  - Unités lexicales font appel à des participants
  - Citron vs. donner;
  - Ordinateur vs. télécharger
- Structure argumentale

```
Ex. vert clou enseigner
```

Un nombre croissant de ressources présente la structure argumentale de manière explicite

## Structure argumentale dans les ressources lexicales (1)

Ressource	Énoncé contenant les participants obligatoires	Nom donné aux participants obligatoires	Étiquettes	Exemple
<u>WordNet</u>	Sentence Frame (pour verbes seulement)			Give
	n the abstract sense or physical se Somebodys somebody somet		ack eye"; "The draft gave	me a cold"
<u>Dicouèbe</u>	Forme propositionnelle; structure actancielle	Actants sémantiques	Variables (X, Y, Z) + étiquette sémantique	Apprécier
apprécier 1 : (act	ion psychique) individu X ~ ; appr	écier 2 : opinion) : individ	u X ~ Y	•
FrameNet*	Frame	Core Frame Element	Étiquettes souvent associées à un seul Frame	Give

#### Frame Giving

A Donor transfers a Theme from a Donor to a Recipient. This frame includes only actions that are initiated by the Donor (the one that starts out owning the Theme). Sentences (even metaphorical ones) must meet the following entailments: the Donor first has possession of the Theme. Following the transfer the Donor no longer has the Theme and the Recipient does.

Donor: The person that begins in possession of the Theme and causes it to be in the possession of the Recipient.

Theme: The object that changes ownership.

Recipient: The entity that ends up in possession of the Theme.



## Structure argumentale dans les ressources lexicales (2)

- Lien entre sémantique et syntaxe
  - WordNet : Une seule représentation (verbes)
  - Dicouèbe : niveau sémantique (variables), syntaxique profond (chiffres romains); syntaxique de surface (régimes)
    - Exemple : apprécier
  - FrameNet : liste des frame elements (niveau conceptuel plutôt que sémantique); annotations et entrées lexicales



## Structure argumentale dans les ressources terminologiques

- connecter
- compatible
- déboisement
- reboisement



### Synonymie

- Synonymie exacte (sans prise en compte des registres ou variantes régionales)
  - Bicyclette, vélo
  - Syn dans Dicouèbe
- Quasi-synonymie
  - Qsyn dans Dicouèbe
- Synsets WordNet (synonymes et quasi-synonymes)



### Antonymie

- compatible; incompatible
- grand; petit
- faire; défaire
- augmenter; réduire
- envoyer; recevoir, acheter; vendre

#### Antonymes dans Dicouèbe

Anti, Conv, Contr



## Hyperonymie

- Relation d'inclusion
- Le sens de l'UL1 est compris dans celui de l'UL2
- Est un, est un type, un X et autres Y
- Hiérarchie
  - Le chat est un félin, le chat et autres félins
  - L'imprimante est un périphérique
- Transitivité
  - Le chat est un félin
  - Le chat et un mammifère
  - Le chat est un animal



### Méronymie

- Relation dans l'espace (lexicalement plus vague que les précédentes)
- Fait partie de, est une partie de
- Hiérarchie
  - Le moteur fait partie de la voiture
- Transitivité non systématique
  - La maison a une porte
  - La porte a une poignée
  - ? La maison a une poignée



## Méronymie

- moteur; voiture
- Meuble; mobilier
- bois; meuble
- Tranche; pain
- boutique; centre commercial
- évaporation; cycle de l'eau



- Relations lexicales dans les ressources
- WordNet
  - Synonymie (synsets)
  - Hyperonymie
    - Cat @ -> feline @ -> carnivore @ -> placental @ -> mammal @ -> vertebrate @ -> chordate @ -> animal @ -> organism @ -> living thing @ -> whole @ -> object @ -> physical entity @ -> entity
  - Lexique complètement structuré
    - Unique beginners <u>cat</u>
  - Méronymie
    - Bicycle
  - Antonymie (adjectifs)



Relations lexicales dans les ressources

#### FrameNet

- Unités lexicales évoquant des cadres sémantiques dans FrameNet (peut contenir des synonymes et quasi-synonymes – et même des antonymes) Change\_of\_Temperature
- Liens entre frames (et non entre unités lexicales)
  - Inheritance (ex. Artifact, Cause\_harm)
  - Precedes, Is preceded by (ex. Committing\_crime)



- Relations lexicales dans les ressources
- Dicouèbe
  - Liens paradigmatiques
    - Synonymie, quasi-synonymie, antonymie
    - Dérivation syntaxique
    - Actancielle; circonstancielle
    - Hyperonymie
    - Méronymie?
  - Liens syntagmatiques
    - Collocations

Exemple vaisselle



- Relations lexicales dans les ressources
- Dicouèbe
  - Fonction lexicale

$$F(x) = y$$

Fonction(mot clé) = valeur de la fonction lexicale

Magn(fumeur) = invétéré

Vaisselle

Real + vaisselle



# Liens lexicaux dans les ressources terminologiques

#### **Autres exemples**

Liens lexicaux

configurer 1; barre d'espacement 1 écosystème

Visualisation graphique



# Références

Cruse, D.A. (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.

DiCoEnviro. Dictionnaire fondamental de l'environnement (<a href="http://olst.ling.umontreal.ca/cgi-bin/dicoenviro/search\_enviro.cgi">http://olst.ling.umontreal.ca/cgi-bin/dicoenviro/search\_enviro.cgi</a>

DiCoInfo. Dictionnaire fondamental de l'informatique et de l'Internet. (<a href="http://olst.ling.umontreal.ca/dicoinfo/">http://olst.ling.umontreal.ca/dicoinfo/</a>

DiCouèbe. Dictionnaire en ligne de combinatoire du français (<a href="http://olst.ling.umontreal.ca/dicouebe/">http://olst.ling.umontreal.ca/dicouebe/</a>)

Fellbaum. C. (Ed.). (1998). WordNet: An Electronic Lexical Database, Cambridge: MIT Press

Fontenelle, T. (Ed.). (2003). FrameNet and Frame Semantics. Special issue of the International Journal of Lexicography 16(3).

FrameNet (https://framenet.icsi.berkeley.edu/fndrupal/)



# Références

Kleiber, G. (1999). *Problèmes de sémantique : la polysémie en question*, Villeneuve d'Ascq (Nord) : Presses universitaires du Septentrion

L'Homme, M.C. (2004). *La terminologie : principes et techniques*. Montréal : Les Presses de l'Université de Montréal.

Mel'čuk, I., Clas A., Polguère A. (1995). *Introduction à La Lexicologie Explicative et Combinatoire*. Louvain-la-Neuve: Duculot.

Polguère, A. (2008). *Lexicologie et sémantique lexicale : notions fondamentales*, 2<sup>e</sup> édition, Montréal: Les Presses de l'Université de Montréal.

Wordnet. A Lexical Database for English (<a href="http://wordnet.princeton.edu/">http://wordnet.princeton.edu/</a>)