

Cours sémantique lexicale 2

Françoise Gayral



**Un exemple de construction de dictionnaire
basée sur la lexicologie explicative et
combinatoire de Melcuk**

<http://www.fas.umontreal.ca/ling/olst/index.html>

**Élaboré au GRESLET (Groupe de Recherche en
Sémantique Et Terminologie) de Montréal (Igor A.
Mel'Cuk, André Clas, Alain Polguère, Louvain la
Neuve), Sylvain Kahane, Paris 7**

La théorie sens-texte



- Théorie linguistique née en Russie dans les années 60 : Zholkovsky & Mel'čuk, maintenant surtout développée par Mel'čuk et ses collègues à Montréal et Paris
- La théorie sens-texte fournit les informations nécessaires pour passer d'une idée (le sens) à sa réalisation dans une langue donnée (le texte)
- La lexicologie explicative et combinatoire est la partie lexicale de la théorie sens-texte.

La théorie sens-texte et le DEC

- Pour Melcuk : le lexique d'une langue prime sur sa grammaire
 - Contraire au point de vue dominant de la linguistique où la syntaxe est reine
- Le DEC est donc central dans le modèle :
 - « Le lexique est au cœur de la linguistique théorique »
- La lexicologie explicative et combinatoire décrit une méthode de construction d'articles de dictionnaire basée sur cette théorie.
 - Cette méthode est indépendante des langues.
 - Elle permet donc de représenter n'importe quelle langue.

Le DEC : un dictionnaire formel



- Un dictionnaire formel, avec une méthodologie de construction précise, utilisant un métalangage complexe
- **Explicatif** : le sens y est décomposé selon des principes rigoureux
- **Combinatoire** : la combinatoire syntaxique et lexicale y est exhaustivement décrite
 - La combinatoire est l'aspect le plus original de l'entreprise.

Une entreprise essentiellement théorique et expérimentale

- **Le DEC = « un schéma idéalisé de dictionnaire de langue »** qui ne prend pas en compte les questions matérielles de temps et de coût
- Le dictionnaire « papier » ne contient que quelques centaines d'articles.
- Le DEC est destiné à un public éclairé
- Le DEC a inspiré des adaptations grand public, comme le *Lexique Actif du Français*
- Le DEC connaît des adaptations pour le TAL, principalement le DiCo (développé par Polguère)

Le DEC = étude des lexies d'une langue



- Une lexie est :
 - un mot pris dans une acceptation spécifique (lexème)
 - une locution pris dans une acceptation spécifique (phrasème)
- une entité à 3 facettes comprenant
 - un sens
 - une forme phonique et graphique
 - un ensemble de traits de combinatoire :
 - genre
 - distribution
 - syntaxique

==> Le DEC n'est pas un dictionnaire de mots

Mot et lexème

mot *pont* -> 15 lexies

- **Lexie 1** : PONT au sens de 'construction reliant les deux rives d'une étendue d'eau'
- **Lexie 2** : PONT au sens de 'jours chomés entre deux jours fériés'
- **Lexie 3** : PONT au sens de 'ensemble des organes transmettant le mouvement'
- **Lexie 4** : PONT au sens de 'circuit électrique formé par 4 composants'
- **Lexie 5** : PONT au sens de 'plancher fermant par en haut la coque d'un bateau'
- ...

Autres lexies (comprenant le mot *pont*)

- Seront des entrées à part entière du DEC plutôt qu'être incluses sous les entrées PONT ou CHAUSSEES ou ...
 - **Lexie 6** : PONT ET CHAUSSÉES au sens de 'service administratif des voies publiques'
 - **Lexie 7** : PONT AÉRIEN au sens de 'liaison régulière par avion entre deux ponts'
 - **Lexie 8** : COUPER LES PONTS au sens de 'cesser délibérément les relations [avec N]'
 - ...

Principes d'élaboration du DEC

- L'élaboration du DEC suit un protocole et des principes formels
- L'activité lexicographique commerciale répond rarement à des principes aussi explicites
- 3 principes de base :
 - Principes de rédaction (Comment ? rédiger une entrée lexicale ?)
 - Principes de nomenclature (Quoi ? Quel est le langage couvert ?)
 - Principes de documentation (A partir de quoi ?)
- DEC : système bien organisé sur 2 plans
 - Microstructure : structure commune à tous les articles du DEC
 - Macrostructure : organisation de l'ensemble des articles du dictionnaire

Principes de rédaction des entrées lexicales



- 4 principes rédactionnels :
 - Principe de formalité
 - Principe de cohérence
 - Principe d'uniformité
 - Principe d'exhaustivité

- formel + cohérent + uniforme + exhaustif = Rigoureux

Principes de rédaction des entrées lexicales

1. Principe de formalité

- D'habitude, descriptions informelles de la représentation sémantique
- Au contraire, le DEC utilise un **métalangage formel préétabli** suffisamment expressif pour :
 - la description sémantique (*définition*)
 - La description syntaxique (*régime*)
 - La description cooccurrentielle (*fonctions lexicales*).
- Les descriptions lexicographiques sont **explicites**. Elles n'omettent rien.

Principe de rédaction des entrées lexicales

1. Principe de formalité

Exemple : MAGAZINE

- **Definition 1**: publication périodique généralement illustrée.
→ **Problème 1** : Ne distingue pas les magazines, des revues ou des journaux illustrés
 - journal : feuilles pas reliées
 - revue : pas forcément pour divertir (*magazine mathématique !*)

- **Definition 2** : publication périodique, sous forme d'un cahier, généralement illustrée, destinée avant tout à divertir.
 - Généralement illustrée : car pas forcément (*magazine de mots croisés...*)

Principe de rédaction des entrées lexicales

1. Principe de formalité

Exemple : MAGAZINE

→ **Problème 2** : N'indique pas que :

- “magazine féminin” veut dire soit “magazine traitant des femmes” (“magazine sportif”) soit “magazine pour les femmes” ;
- “magazine enfantin” ne peut se dire (il faut dire “magazine pour les enfants”), ...
- **Definition 3** : publication périodique, sous forme d'un cahier, généralement illustrée (consacrée à un sujet X et) destinée avant tout à divertir le public Y.
- Pour les variables X et Y, le **régime** (une sous-zone spéciale de l'article de dictionnaire) spécifiera la façon de les exprimer :
 - | X = politique, sportif, de théâtre (et pas théâtral), féminin, de cinéma
 - | Y = féminin/pour femmes, pour les enfants etc.

Formel = écrit dans un métalangage préétabli + explicite

Principes de rédaction des entrées lexicales

2. Principe de cohérence

- Un article de dictionnaire est cohérent s'il met en correspondance
 - les composants sémantiques dans la définition
 - les actants syntaxiques profonds
 - ses cooccurrents lexicaux
- Des variables font le lien entre lexème et actants syntaxiques dans la phrase
- Pour chaque actant syntaxique, le **régime** spécifie l'expression correcte à la surface (c-à-d dans la réalisation concrète de la phrase)
- Exemple : X et Y pour MAGAZINE

Principes de rédaction des entrées lexicales

2. Principe de cohérence

Exemple : BLESSURE

- **Première définition** : “lésion visible faite aux tissus de l’être vivant par une cause extérieure”
 - **Problème** : Ne permet pas de rendre compte de la cooccurrence avec “grave, mortelle, sérieuse, légère”
- **Définition DEC** : “lésion visible faite aux tissus de l’être vivant par une cause extérieure et qui porte atteinte à la santé de X”
 - explicite la composante **intensifiable**
 - permet de rendre compte de façon systématique des cooccurrents possibles grâce à la fonction lexicale standard de 'Magnitude'

Cohérent = tel que la sémantique, la syntaxe et la cooccurrence lexicale sont mis en correspondance

Principes de rédaction des entrées lexicales

3. Principe de traitement uniforme

- Les descriptions de lexies sémantiquement liées doivent avoir la même structure

- Exemple :

- revue, magazine, journal...
- français/chinois/breton/espagnol etc.

ébauche de description : “ethnie originaire du pays X dont la langue maternelle est le Y”

mais pb des états multi-nationaux, des ethnies sur plusieurs pays (les berbères), ethnie ?? (les Corses), langue maternelle ??

- Exigence de cohérence appliquée à la macrostructure du dictionnaire

Uniforme = traitement similaire pour toutes les unités d'un champ sémantique

Principes de rédaction des entrées lexicales

4. Principe d'exhaustivité

- **externe** : toutes les lexies
- **interne** : Pour chaque lexie, toutes les informations lexicographiques nécessaires
 - pour bien utiliser cette lexie
 - pour trouver les lexies qui lui sont sémantiquement liées

Exemple : L'entrée pour le phrasème **Transport en commun** inclut :

- les noms des divers moyens de transport
- des lexies comme **titre de transport, voyageur, correspondance, réseau, desservir, transporter, voyager, composer** etc.

Exhaustif = totalité des informations pour chaque lexie, y compris les références à toutes les lexies apparentées

Principes de nomenclature

Quel est le langage couvert ?

- Nomenclature = ensemble des lexies décrits
- La nomenclature du DEC s'appuie sur 2 principes :
 - le principe de la disponibilité
 - le principe de lexicalité
- Principe de disponibilité : la nomenclature du DEC ne comprend que le **lexique disponible** c-à-d appartenant à la **langue générale standard contemporaine**
 - générale = usage linguistique des gens ayant atteint un certain niveau de scolarité : axe sociolectal (classes et groupes sociaux)
 - standard = français hexagonal : axe topolectal (lieu)
 - contemporaine = 50 dernières années, axe chronolectal (temps)

Principes de nomenclature

quelles sont les entrées du DEC ?

- Principe de lexicalité = le DEC décrit toutes les lexies de la langue, lexèmes (**mots simples**) et **phrasèmes**
Centre commercial, arriver comme un cheveu sur la soupe, dos d'âne, tomber dans les pommes, mort de fatigue
- Typologie fine des phrasèmes
 - repose sur des critères sémantiques
 - ne conduit pas à un traitement uniforme dans le DEC

Entrées du DEC : les phrasèmes

Typologie des phrasèmes

- **Phrasèmes complets** (figés, expressions idiomatiques) : le sens du tout n'est pas déductible du sens des parties
Exemple : *pomme de terre, œil de bœuf, cordon bleu, tomber dans les pommes*
- **Quasi-phrasèmes** : le sens du phrasème inclut le sens de ses constituents plus un surplus de sens imprévisible
Exemple : *carte bleue, donner le sein, centre commercial*
- **Phrasèmes complets et quasi-phrasèmes** forment des entrées indépendantes

Typologie des phrasèmes

- **semi-phrasèmes ou collocations :**
phrasème dans lequel l'un des constituants (la base) conserve son sens alors que l'autre (le collocatif) soit ne garde pas son sens, soit n'est pas sélectionné librement
- Exemple : *piquer un somme, accepter/décliner une invitation, peur bleue, fort comme un turc, mort de fatigue*
- Les semi-phrasèmes sont codés à l'intérieur des articles de base de leurs constituants, à l'aide de **fonctions lexicales**

Entrées du DEC

liens phrasèmes/lexèmes

- Pour faciliter l'accès du lecteur aux phrasèmes, le DEC utilise des renvois sémantiques et formels :
 - Renvois *sémantiques* entre toutes les lexies (mots et phrasèmes)
Exemple : **fusil/viser/tirer** renvoie à “**coucher en joue**”, **secret** renvoie à “**découvrir le pot aux roses**”
 - Renvois *formels* à un phrasème sous tous ses constituants
Exemple : “**coucher en joue**” est référencié sous “**coucher**” et sous “**joue**”

Après ces précautions méthodologiques, on y va...

Microstructure des articles

- Avant tout : Comment le DEC répond à la question polysémie/homonymie ? Quels critères de délimitation des lexies ?
- Critères purement linguistiques
 - Critère de dérivation : lexies différentes si pas même dérivation
 - | *Jeanne élève(1) des cochons/ Jeanne élève(2) ses cochons*
 - | *élève(1) → élevage, éleveur...*
 - | *Blanchir le linge, blanchir l'argent*
 - Critère d'interprétation multiple : lexies différentes s'il existe une phrase qui a 2 interprétations possibles à cause de L
C'est Jean qui a peint le plafond

Critères de délimitation des lexies

- Critère de cooccurrence compatible : pas de lexies différentes s'il existe une phrase qui autorise les 2 sens

? *Jean a peint trois tableaux et le mur de sa maison*

? *Jean s'occupe de ses enfants et de linguistique*

Jean s'occupe de ses enfants et de sa maison

- Critères de co-occurrence différentielle: lexies différentes si on peut :

- dégager des constructions syntaxique différentes

- Exemple *changer* intransitif

changer de direction/de chemise/ de téléphone vs changer de chambre/d'école

==> Existence d'une tournure transitive et causative

*Changer Jean d'école vs *Changer Jean de chemise/ de téléphone*

- dégager des cooccurrents lexicaux différents

blessé (physiquement) et blessé (moralement)

Intensité : grièvement F. Gay et al profondément

Structure de l'article

Description de chaque lexie selon plusieurs axes :

- son sens (cf. définition)

→ Zone sémantique

- sa combinatoire syntaxique (cf. régime)
- ses dérivés sémantiques (fonctions lexicales paradigmaticques), les collocations qu'elle contrôle (fonctions lexicales syntagmaticques)

→ Zone combinatoire

- des illustrations de son comportement en langue (cf. exemples)

Zone sémantique



- Décrit le contenu sémantique de la lexie
- Comprend deux zones :
 - Définition lexicographique
 - Décomposition du sens de l'entrée lexicale en sens plus simples
 - Fait apparaître les actants sémantiques
 - Se fait principalement par **paraphrasage**
 - Connotations (parfois absentes)

Définition lexicographique

exemples

- Sous forme rédigée

Défini

Professeur de Y/de Z (à W) = ‘ Personne qui, en tant que profession, **enseigne** Y d ’un niveau assez élevé à Z (étant employée par un établissement d ’enseignement W)’

Définissant
(definiens)

() : facultatif

X enseigne Y à Z = ‘ X cause volontairement que A **apprenne** Y, en transmettant systématiquement à Z des connaissances (portant) sur Y ou des techniques (portant) sur Y ‘

Actant sémantique

- La notion d'actant sémantique (argument d'un prédicat) est essentielle : ASém
- **Prédicat** : sens qui présuppose des « participants » sémantiques
- **Actant sémantique** : participant associé au prédicat (argument du prédicat)
 - Prédicat unaire (un actant sém.) : ‘ gentil ’, ‘ ronfler ’, ‘ tête ’ ...
 - Prédicat binaire (deux actants sém.) : ‘ lire ’, ‘ tristesse ’, ‘ fier ’ ...
 - Prédicat ternaire (trois actants sém.) : ‘ donner ’, ‘ obtenir ’ ..

Actant sémantique

- Toutes les lexies ne sont pas des prédicats : EAU, FLEUR, SOLEIL...
- La plupart des lexies sont des prédicats et ont donc des ASém :
 - Verbes, adjectifs, adverbes, conjonctions, prépositions => toujours!
 - Noms => souvent (même pour des noms concrets, comme *nez*, *écran*, *toit* ...)
 - Jusqu'à 6 Asem : Exiler "X exile Y de L1 à L2 pour Z pendant la période T"
- Les ASém ne sont pas tous réalisés en « surface » dans une phrase où L apparaît
 - Ex : *Pierre a vendu sa voiture*
Il n'y a ici ni destinataire, ni somme d'argent (mais ces deux actants sont sous-jacents).

==> Une notion très large des prédicats !

Exercice



- Trouver les actants sémantiques de LETTRE, INTERDIRE, FACTURE

Règles d'élaboration de la définition

- Pour une lexie prédicative L, les actants sémantiques de L doivent apparaître dans le défini et dans le définissant
- Une lexie doit être définie avec des lexies plus simples pour éviter la circularité des définitions
 - *Astronomie* + simple que *astronome*
Astronomie: science des corps célestes *astronome* personne qui en tant que profession fait de l'astronomie
- Eviter les décompositions trop fines
- Eviter les termes ambigus (donc préciser l'acception du mot en utilisant son numéro) et les termes synonymes dans le définissant
 - Montre* “dispositif pour lire l'heure” *marteau* “outil de percussion”
 - couteau* “instrument servant à couper”
 - cuillère* “ustensile qui sert à porter la nourriture à la bouche”
 - 4 formes : ***pour/de/ servant à/qui sert à pour dire la même chose*** donc choix d'une seule expression ***destiné à***

Quelles composantes mettre dans la définition ?

- Celle qui permet de faire le lien entre différents sens de la même lexie
 - composante ‘blanc’ pour la lexie *neige* car :
 - | Existence de *blanc comme neige*
 - | Existence d’autres lexies *neige* (cocaïne) liées par cette composante
 - Idem pour *sel* (à cause de *poivre et sel*) mais pas pour *riz*
 - Celle qui reflète certaines cooccurrences (avec des adjectifs)
 - Exemples : On l’a vu avec *blessure*
 - **Applaudissement** : modifiable par des intensificateurs/atténuateurs (nourris, frénétiques, clairsemés etc.)
- ==> la composante d’intensification doit apparaître dans la définition
- **Applaudissement de X à Y pour Z** = “Battements de mains par ^{F. Gayral} **Z**”

Deuxième partie de la zone sémantique : Connotations

- La sémantique d'une lexie n'est pas toujours couverte par sa définition
- Les connotations rendent compte des usages figurés de la lexie en explicitant les liens sémantiques entre la lexie et des expressions figurées contenant la lexie

Vent définition : VENT de X= 'phénomène atmosphérique- déplacement horizontal perceptible de l'air¹ venant de X, X étant un lieu géographique ou un des points cardinaux'

- Vent connotations
 - 1) ' rapidité ' => *en coup de vent*
 - 2) ' changement d'orientation ' => *le vent tourne*
 - 3) ' force motrice ' => *avoir le vent en poupe*
 - 4) ' absence de contenu ' => *c'est du vent.*

Zone de combinatoire



- La Zone de combinatoire se divise en :
 - **zone de combinatoire syntaxique** _____
 - **zone de combinatoire lexicale restreinte** _____
 - zone de combinatoire stylistique (type de contexte textuel dans lequel la lexie peut s'insérer e.g., litt., fam., arch. etc.)
 - zone de combinatoire morphologique (partie du discours, type de déclinaison, formes irrégulières etc.)

Zone de combinatoire syntaxique

- Un actant syntaxique profond de la lexie L est un syntagme qui dépend de L syntaxiquement et qui en exprime un actant sémantique
 - Notation : numérotés avec des chiffres romains
- Un actant syntaxique profond (ASP) peut correspondre à un complément circonstanciel
 - **X se trouver à Y** : le complément circonstanciel de lieu = ASP II exprimant l'actant sémantique Y
- ASP pas seulement pour les verbes mais aussi pour les noms
 - **Arrivée de X** (*arrivée de Pierre*): Pierre = ASP I de arrivée

Zone de combinatoire syntaxique = tableau de régime

- Tableau de régime d'une lexie :
 - Colonnes : les actants sémantiques
 - Lignes : possibilités d'expression pour les actants syntaxiques profonds
- Exemple

X vendre I.1 Y à Z pour W

Paul a vendu sa voiture à Pierre pour 2000 euros

X=I	Y=II	Z=III	W=IV
1. N	1. N obligatoire	1. à N	1. Num N 2. à num N 3. pour N

Tableau de régime

- DESESPOIR 1 = Désespoir de X devant Y

X = I	Y = II
<p>1. <i>de</i> N 2. <i>Aposs</i> 3. <i>A</i></p>	<p>1. <i>de</i> N 2. <i>à</i> N 3. <i>devant</i> N 4. <i>de</i> Vinf</p>

Actant sémantique

Actant syntaxique

Réalisation de surface des actants syntaxiques profonds

Zone de combinatoire lexicale

- Cette zone vise à :
 - recenser les possibilités d'emplois d'une lexie dans un même **paradigme** sémantique : liens aux synonymes, antonymes, converses,...
 - guider le choix des cooccurrents lexicalement contraints : collocations, contraintes **syntagmatiques** lexicales, ...
 - a. *Fort comme un Turc ; Jaloux comme un tigre*
 - b. ?? *Fort comme un tigre ; Jaloux comme un Turc*
- Cette zone permet donc le choix du mot juste (en opposition avec d'autres mots possibles) et la production de la combinaison adéquate
- Cette zone est spécifiée à l'aide de **fonctions lexicales** (52), indépendantes des langues

Zone de combinatoire lexicale

- Les fonctions lexicales (FL) sont de la forme $f(x) = y$
 - x : une lexie (mot-clé)
 - y : ensemble de lexies
- Exemple : la fonction **Magn** exprime l'intensité
 - Magn(*fort*) = {*comme un turc*}
 - Magn(*peur*) = {*bleue, grande, immense ...*}
 - Magn(*travailler*) = {*d'arrache-pied, énormément, comme un fou*}
 - Magn(*pleurer*) = {*amèrement, à chaude larmes, comme une madeleine, ...*}
 - Magn(*pluie*) = {*grosse, diluvienne, violente, ...*}

Zone de combinatoire lexicale : fonctions lexicales

- 56 fonctions lexicales couvrent
 - les fonctions **syntagmatiques** qui relient des lexies apparaissant en relation de cooccurrence : verbe support, modifieurs spécifiques, ...
 - ┆ Exemple de FL syntagmatique
 - Oper₁(visite) = rendre
 - Oper₂(visite) = recevoir
 - Mult(soldat) = régiment
 - les fonctions **paradigmatiques** qui relient des lexies n'apparaissant pas en cooccurrence immédiate et qui regroupent toutes les relations sémantiques de type synonymie, antonymie, hyperonymie...
 - ┆ Exemple de FL paradigmatique
 - Syn(voiture) = automobile
 - Gener(cerise) = fruit

Combinatoire lexicale : FL paradigmatiques

- **Synonyme [Syn]** (4 types de synonymie traités)
 - Syn (voiture) = auto(mobile)
- **Conversif [Conv]** : mêmes actants sémantiques, actants syntaxiques profonds inversés
 - Conv (craindre) = effrayer
- **Antonyme [Anti]**
 - Anti (respect) = irrespect
 - AntiMagn(majorité) = faible, courte
- **Contrastif [Contr]**
 - Contr (glace) = feu
 - Contr (ciel) = terre
- **Générique [Gener]**
 - Gener (armoire) = meuble
 - Gener(*tristesse*) = *sentiment*
- **Figuratif [Figur]**
 - Figur(fumée) = rideau de ~

Combinatoire lexicale

FL paradigmatiques

- Dérivé syntaxique : nominalisation [S0], verbalisation [V0],
adjectivisation [A0], adverbialisation [Adv0]
 - S0 (acheter) = achat V0 (sermon) = jurer
- Dérivé sémantique nominal actanciel [Si]
 - S1 ~ agent S1 (parler) = locuteur
 - S2 ~ patient S2 (parler) = parole, propos...
 - S3 ~ but/ bénéficiaire S3 (parler) = destinataire ...
- Dérivé sémantique adjectival potentiel [Able] (souvent combiné avec Anti)
 - Able (lire) = lisible AntiAble (lire) = illisible
- Dérivations circonstancielles
 - S_{loc} (patiner) = patinoire S_{instr} (peindre) = pinceau

Combinatoire lexicale

FL syntagmatiques

- FL adjectivales, ex. : Intensificateur [Magn], Laudatif [Bon] (souvent combiné avec Anti)
 - | Magn (boire) = comme un trou
 - | AntiBon (temps) = de chien
- FL adverbiales, ex. : Instrumental [Instr]
 - | Instr (outil) = avec [ART ~] Instr (téléphone) = par [~]
- FL verbales, ex. : verbes supports (sémantiquement vides en cooccurrence avec des noms prédicatifs) [Oper_i], verbes de réalisation [Real_i ou Fact_i]
 - Oper₁(visite) = rendre Oper₂(visite) = recevoir
 - Real₂(hypothèse) = confirmer
 - Fact₁(souhait) = s'accomplir, se réaliser

Bilan



- On a vu :
 - la microstructure du DEC, c'est-à-dire l'organisation interne d'un article de dictionnaire
 - la méthodologie pour construire un tel article
 - le contenu des articles

- sur la base : une lexie ↔ un article

- Mais quel regroupement des lexies ? Quelle organisation le l'ensemble du dictionnaire ?
 - **Macrostructure**

La macrostructure des articles

- Le DEC regroupe les lexies correspondant au même mot et qui sont reliées sémantiquement sous un même **vocable**
→ **Regroupement de articles de polysèmes dans un super-article appelé vocable**

Exemple : Le vocable **DESESPOIR** regroupe deux lexies :

- DESESPOIR 1 : l'émotion
- DESESPOIR 2 : l'objet du désespoir 1

- Les lexies qui n'ont aucun point sémantique commun seront décrites dans des vocables différents

Exemple:

PIED¹I : ' partie du corps ... '

PIED² : ' unité rythmique minimale d'un vers '

Concepts de base pour établir les relations entre lexies dans le vocable

- **Notion de pont sémantique**
une composante sémantique caractéristique commune entre deux lexies.
- **Lien direct/indirect**
 - Deux signifiés sont liés **directement** s'ils ont un pont sémantique.
 - Deux signifiés sont liés **indirectement** si l'on peut établir une chaîne de signifiés intermédiaires ayant des ponts sémantiques.

Notion de pont sémantique : lien direct ou indirect

■ Exemple

- MINE I : ‘ cavité creusée sous la terre destinée à extraire un minerai. ’
- MINE II.1 : ‘ cavité creusée sous la terre, située sous une fortification, contenant une charge explosive et destinée à faire sauter cette fortification. ’
- MINE II.2 : ‘ engin contenant une charge explosive et destiné à faire sauter ... ’



Distances sémantiques entre les lexies

- (classiquement) marquées par des numéraux lexicaux
 - **Grandes distances : chiffres romains**
 - **Moyennes distances : chiffres arabes**
 - **Petites distances : lettres**

Vocable : Adhésion

I. événement physique ~de substance X à substance Y

II.1. action mentale ~de personne X à idée Y

II.2.a. action sociale ~de personne X à organisation Y

2.b. document ~de personne X [à organisation Y]

\$

Organisation dans le champ lexical

- L'étude des lexies se fait par champ lexical : principe de traitement uniforme

- Mais définir le champ lexical n'est pas évident !
 - plusieurs points de vue sur les lexies :
 - | *chien* : champ des mammifères ou animal de compagnie ??
 - | *œuf* : champ des comestibles ou objet transportable ??
 - | *Avion* : champ transport ou objet construit (artefact) mécanique

Applications du DEC

- Le DEC sous sa forme traditionnelle papier :
 - est trop complexe.
 - n'est pas suffisamment formalisé, même si le modèle a d'emblée été conçu dans la perspective d'une informatisation.

- D'où deux dictionnaires “dérivés”
 - **LAF** : Lexique Actif du Français : dictionnaire grand public de dérivations sémantiques et de collocations du français
 - **DiCo** (dictionnaire de cooccurrences) : version simplifiée et informatisée du DEC (~383vocables)
 - centré sur la description des liens lexicaux paradigmatiques (dérivation) et syntagmatiques (collocations) plutôt que sur la description de leur sens.
 - pas de véritable définition lexicale mais un étiquetage sémantique des lexies et la description de leur structure actancielle.

Applications du DEC

■ DiCo avec

- un éditeur du DiCo : EdiCo pour :
 - l'édition d'articles de DiCo (le squelette de chaque article est généré automatiquement par ex.)
 - la recherche sur la base du dictionnaire
- une interface permettant de consulter en ligne une version compilée du DiCo : Le DiCouèbe est
 - <http://olst.ling.umontreal.ca/dicouebe/index.php>

■ Applications en didactique

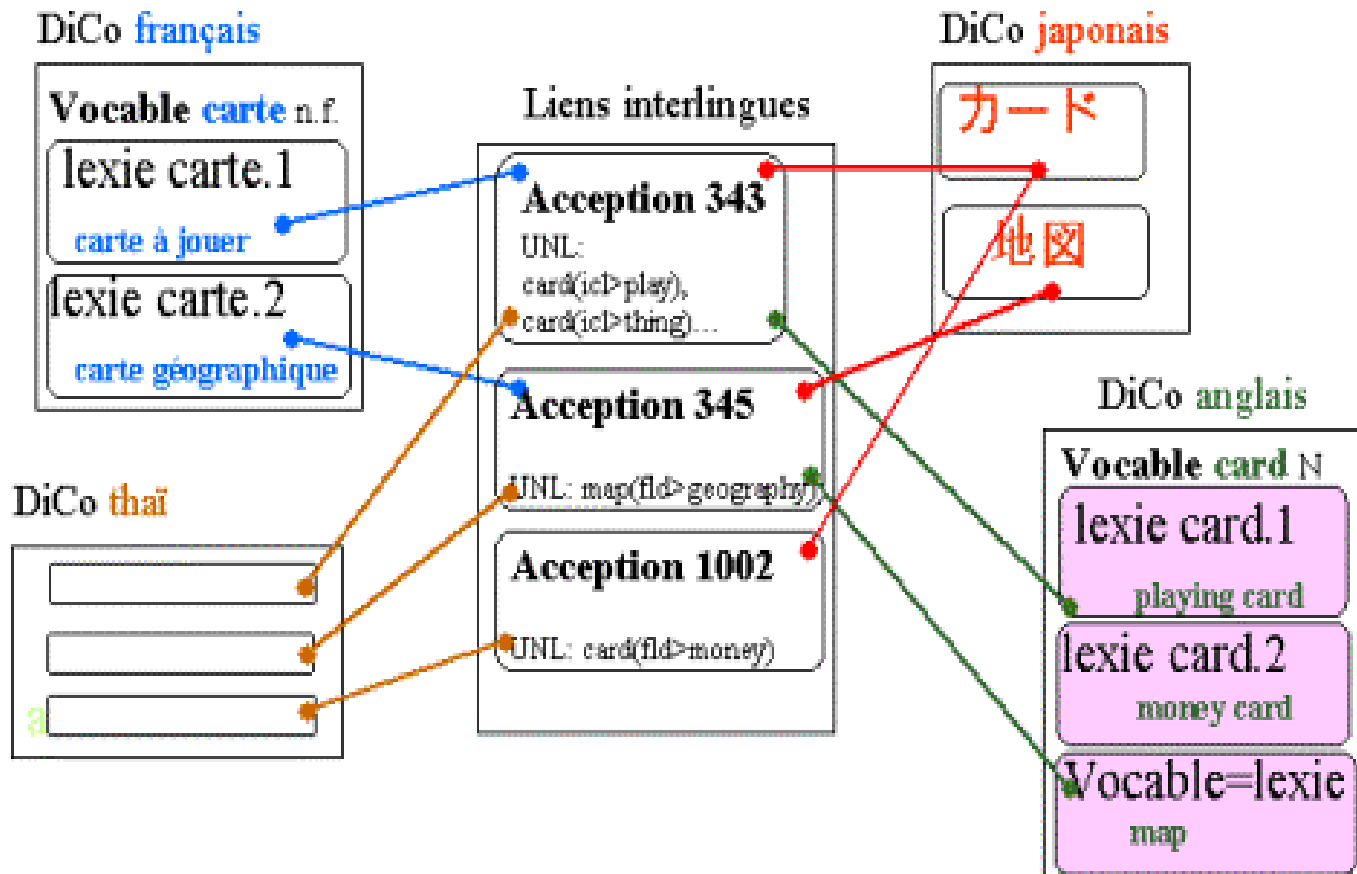
- Enseignement du français langue étrangère
- Aide à la production de textes en primaire et au collège

■ Applications en TAL

- Paraphrase
- Génération de texte : étape de lexicalisation [Polguère]

Le projet Papillon ~2000

- But : construire une base lexicale multilingue (français, japonais, anglais, thaï) disponible sur le web



Le projet Papillon



- les dictionnaires sont basés sur la lexicologie explicative et combinatoire
- L'architecture de la base lexicale est symétrique, "à pivot" : N dictionnaires monolingues accédés par vocables et par lexies (sens de mots) et un ensemble d'axies (liens interlingues)
- Toutes les données de la base lexicale sont décrites sous forme de documents XML
- Les lexies, <lexie> en XML, contiennent 10 éléments principaux : nom du vocable, prononciation, éléments spécifiques aux langues, catégorie grammaticale, formule sémantique, régime, fonctions lexicales, exemples, idiotismes et liens vers les axes.

Exemple de représentation : la lexie meurtre extrait

```
<lexie id="meurtre$1" history="h01" basic="true" >
  <headword hn="1">meurtre</headword>
  <pos>n.m.</pos>
  <semantic-formula>
    action de tuer: ~ PAR L'
    <actor>  <sem-label>individu</sem-label>
              <sem-variable>X</sem-variable>
    </actor>
              DE L'
    <actor>  <sem-label>individu</sem-label>
              <sem-variable>Y</sem-variable>
    </actor>
  </semantic-formula>
```

Les actants
sémantiques

```

<government-pattern>
  <mod nb="1">
    <actor>
      <sem-actant>X</sem-actant>=
      <synt-actant>I</synt-actant>=
      <surface-group>
        <surface>
          <reflexie href="#de">de</reflexie>N
        </surface>,
        <surface>A-poss</surface>
      </surface-group>
    </actor>
    <actor>
      <sem-actant>Y</sem-actant>=
      <synt-actant>II</synt-actant>=
      <surface-group>
        <surface>
          <reflexie href="#de">de</reflexie>N
        </surface>,
        <surface>A-poss</surface>
      </surface-group>
    </actor>
  </mod>
</government-pattern>

```

régime : tueur et victime d'un meurtre peuvent être exprimés par un complément prépositionnel précédé de *de* ou par des adjectifs possessifs.

<lexical-functions>

<function name="V0">

<valgroup>

<reflexie href="#tuer\$1">tuer</reflexie>

</valgroup>

</function>

</lexical-functions>

<examples>

<example id="e1">C'est ici que le double meurtre a été commis.</example>

</examples>

<full-idioms>

<idiom id="i1" href="papillon-axi.xml#i04">_appel au meurtre_</idiom>

</full-idioms>

<axies>

<refaxie href="papillon-axi.xml#a01"/>

</axies>

</lexie>

la combinatoire
lexicale

Conclusion



- Faiblesses de l'approche DEC
 - Lenteur du travail lexicographique
 - Lourdeur de la description (formalismes, conventions d'écriture, etc.)
 - Encodage semi-formel des définitions
 - Difficulté du traitement manuel et du traitement informatique
 - Disparités entre les volumes publiés
 - Absence d'une version électronique complète

Différentes sources de ce cours



- Agnès Tutin (Grenoble)
- Claire Gardent (Nancy)
- Jacques Moeschler (Genève)
- Benoit Habert (Nanterre)
- Pollet Samvelian (Paris 3)
- Henriette Gezundhajt (Toronto)