

Vers un traitement automatique de la néosémie : approche textuelle et statistique

Coralie Reutenauer

coralie.reutenauer@atilf.fr

Soutenance de thèse en Sciences du Langage
20 janvier 2012 — ATILF, Nancy, France

Jury

Rapporteurs

Anne Condamines, Directrice de recherche CNRS, CLLE-ERSS, Toulouse

Jean-François Sablayrolles, Professeur des universités Paris XIII, LDI, Paris

Examineurs

Ludovic Lebart, Directeur de recherche CNRS, Télécom-Paristech, Paris

Alain Polguère, Professeur des universités, Université de Lorraine, ATILF, Nancy

Jean-Marie Pierrel, Professeur des universités, Université de Lorraine, ATILF, Nancy

Evelyne Jacquy, Chargée de recherche CNRS, ATILF, Nancy

Mathieu Valette, Professeur des universités, Inalco, Paris

La langue en évolution

Les habitants de la rue Siabas
à **Outreau** (Pas-de-Calais)...

L'affaire d'**Outreau**, un
véritable fiasco judiciaire

Les **tsunamis** d'origine
volcanique sont plus rares...

alors qu'un « **tsunami bleu** »
s'annonce avec un nombre
record de députés pour l'UMP...

Dettes US : Prémices du
tsunami financier

...émissions de gaz **toxiques**..

...des banques qui ont distribué
des prêts **toxiques**...

...**initier** le grand public...

...**initier** le débat...

...**initier** une conversation...

...**bouclier fiscal**...

...**globalisation**...

La langue en évolution

Les habitants de la rue Siabas
à **Outreau** (Pas-de-Calais)...

L'affaire d'**Outreau**, un
véritable fiasco judiciaire

Les **tsunamis** d'origine
volcanique sont plus rares...

alors qu'un « **tsunami bleu** »
s'annonce avec un nombre
record de députés pour l'UMP...

Dettes US : Prémices du
tsunami financier

...émissions de gaz **toxiques**..

...des banques qui ont distribué
des prêts **toxiques**...

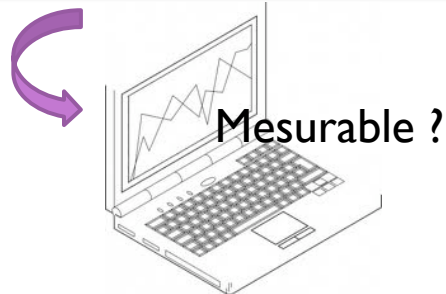
...**initier** le grand public...

...**initier** le débat...

...**initier** une conversation...

...**bouclier fiscal**...

...**globalisation**...



La langue en évolution

Les habitants de la rue Siabas
à **Outreau** (Pas-de-Calais)...

L'affaire d'**Outreau**, un
véritable fiasco judiciaire

Les **tsunamis** d'origine
volcanique sont plus rares...

alors qu'un « **tsunami bleu** »
s'annonce avec un nombre
record de députés pour l'UMP...

Dettes US : Prémices du
tsunami financier

...émissions de gaz **toxiques**..

...des banques qui ont distribué
des prêts **toxiques**...

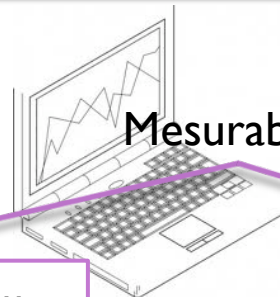
...**initier** le grand public...

...**initier** le débat...

...**initier** une conversation...

...**bouclier fiscal**...

...**globalisation**...



Mesurable ?

Existence d'un nouveau
sens ?

Quel est le nouveau sens ?

Thèse défendue

Il est possible de construire de façon semi-automatique
une représentation structurée et nuancée d'un nouveau sens lexical

Thèse défendue

Quel phénomène
linguistique ?

Nouveauté
par rapport à
quoi ?

Il est possible de construire de façon semi-automatique
une représentation structurée et nuancée d'un nouveau sens lexical

Thèse défendue

Quel phénomène linguistique ?

Nouveauté par rapport à quoi ?

Il est possible de construire de façon semi-automatique une représentation structurée et nuancée d'un nouveau sens lexical

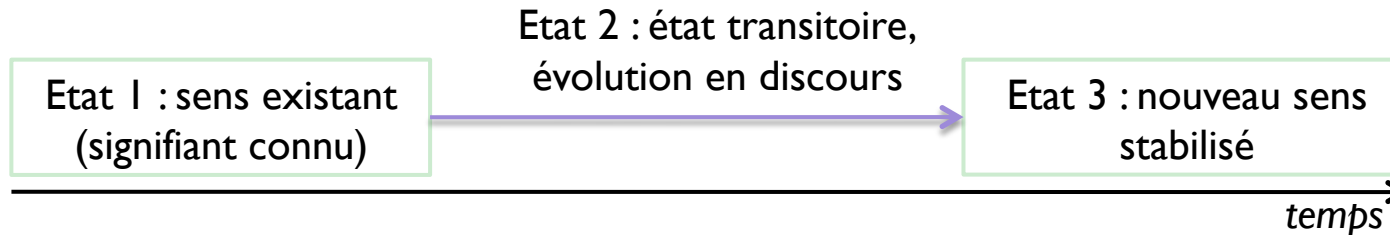
Où chercher l'information ?

Quel modèle de représentation ?

Comment ?

Néologie sémantique et perspective applicative

(Bastuji, 1974) – (Sablayrolles, 2000) – (Mejri, 2010)



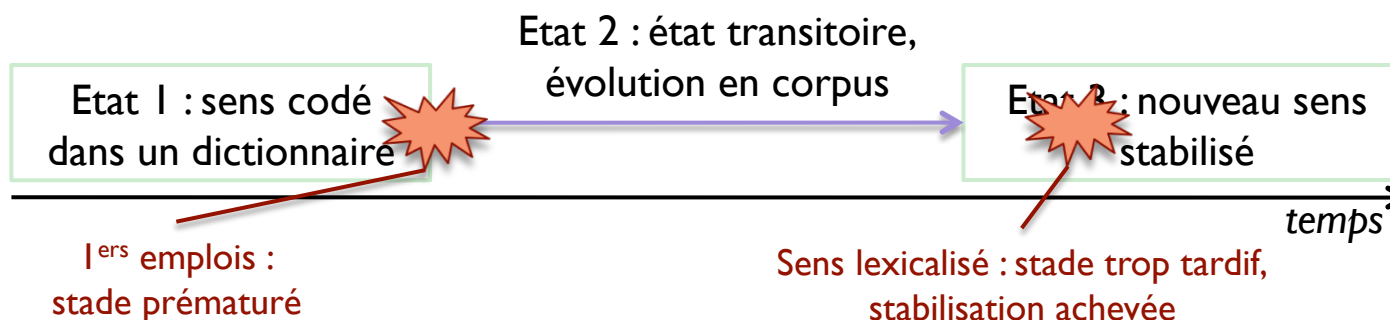
Néologie sémantique

changement de sens caractérisé par l'apparition d'un nouveau signifié pour un signifiant préexistant

- ▶ En linguistique
 - ▶ Processus (stade de diffusion variable)
 - ▶ Écart entre un **sens existant** et de **nouveaux sens en discours**

Néologie sémantique et perspective applicative

(Bastuji, 1974) – (Sablayrolles, 2000) – (Mejri, 2010)



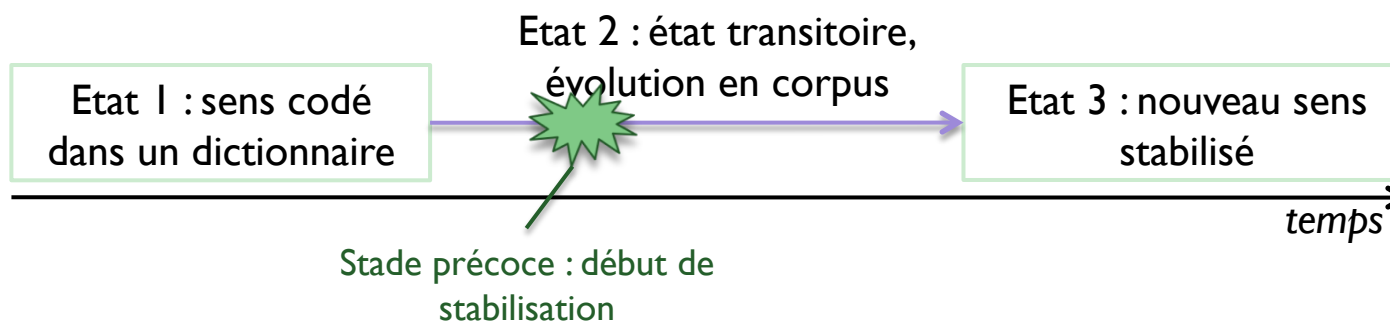
Néologie sémantique

changement de sens caractérisé par l'apparition d'un nouveau signifié pour un signifiant préexistant

- ▶ En linguistique
 - ▶ Processus (stade de diffusion variable)
 - ▶ Écart entre un **sens existant** et de **nouveaux sens en discours**
- ▶ Dans notre approche applicative
 - ▶ Processus : diffusion en cours
 - ▶ Écart entre un **sens codé dans une ressource de référence** (dictionnaire) et des **emplois dans des ressources discursives** (corpus de textes)

Néologie sémantique et perspective applicative

(Bastuji, 1974) – (Sablayrolles, 2000) – (Mejri, 2010)



Néologie sémantique

changement de sens caractérisé par l'apparition d'un nouveau signifié pour un signifiant préexistant

- ▶ En linguistique
 - ▶ Processus (stade de diffusion variable)
 - ▶ Écart entre un **sens existant** et de **nouveaux sens en discours**
- ▶ Dans notre approche applicative
 - ▶ Processus : diffusion en cours
 - ▶ Écart entre un **sens codé dans une ressource de référence** (dictionnaire) et des **emplois dans des ressources discursives** (corpus de textes)

Indices de néosémie : le choix des cooccurrents

« *Tout néologisme de sens ne se repère que par la bizarrerie ou la nouveauté de son environnement* »
(Gilbert, 1974)



Néosémie

« *innovation sémantique dont témoignent les emplois* » (Rastier et Valette, 2009)

➔ Néologie sémantique au prisme des indices délivrés par le contexte linguistique

Indices de néosémie : le choix des cooccurrents

« Tout néologisme de sens ne se repère que par la bizarrerie ou la nouveauté de son environnement »
(Gilbert, 1974)

Néosémie

« innovation sémantique dont témoignent les emplois » (Rastier et Valette, 2009)

→ Néologie sémantique au prisme des indices délivrés par le contexte linguistique

(Bauer et Renouf, 2000) – (Picton, 2009) – (Dury, 2008)

Indices de détection

- guillemets, expressions introductives ou consécutives
- empreintes de fréquence

Indices de caractérisation

- **cooccurrents**

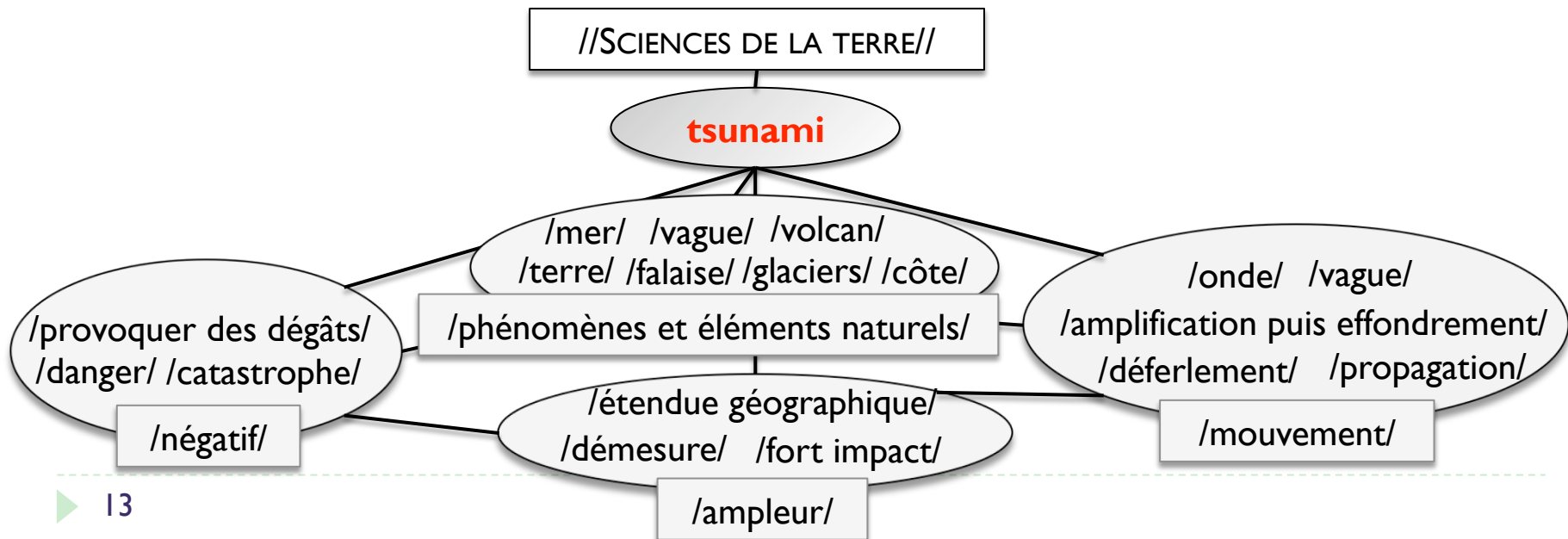
...des *banques* qui ont *distribué* des *prêts toxiques*...

- ▶ Intérêt des cooccurrents (Bastuji, 1974) – (Mayaffre, 2008)
 - ▶ qualification du nouveau sens
 - ▶ favorable à une approche quantitative
 - ▶ appréhension non immédiate

Représentation du sens comme ensemble de sèmes (Rastier *et al.*, 1994) – (Rastier et Valette, 2009)

Cadre théorique
Inspiré de la **sémantique textuelle**

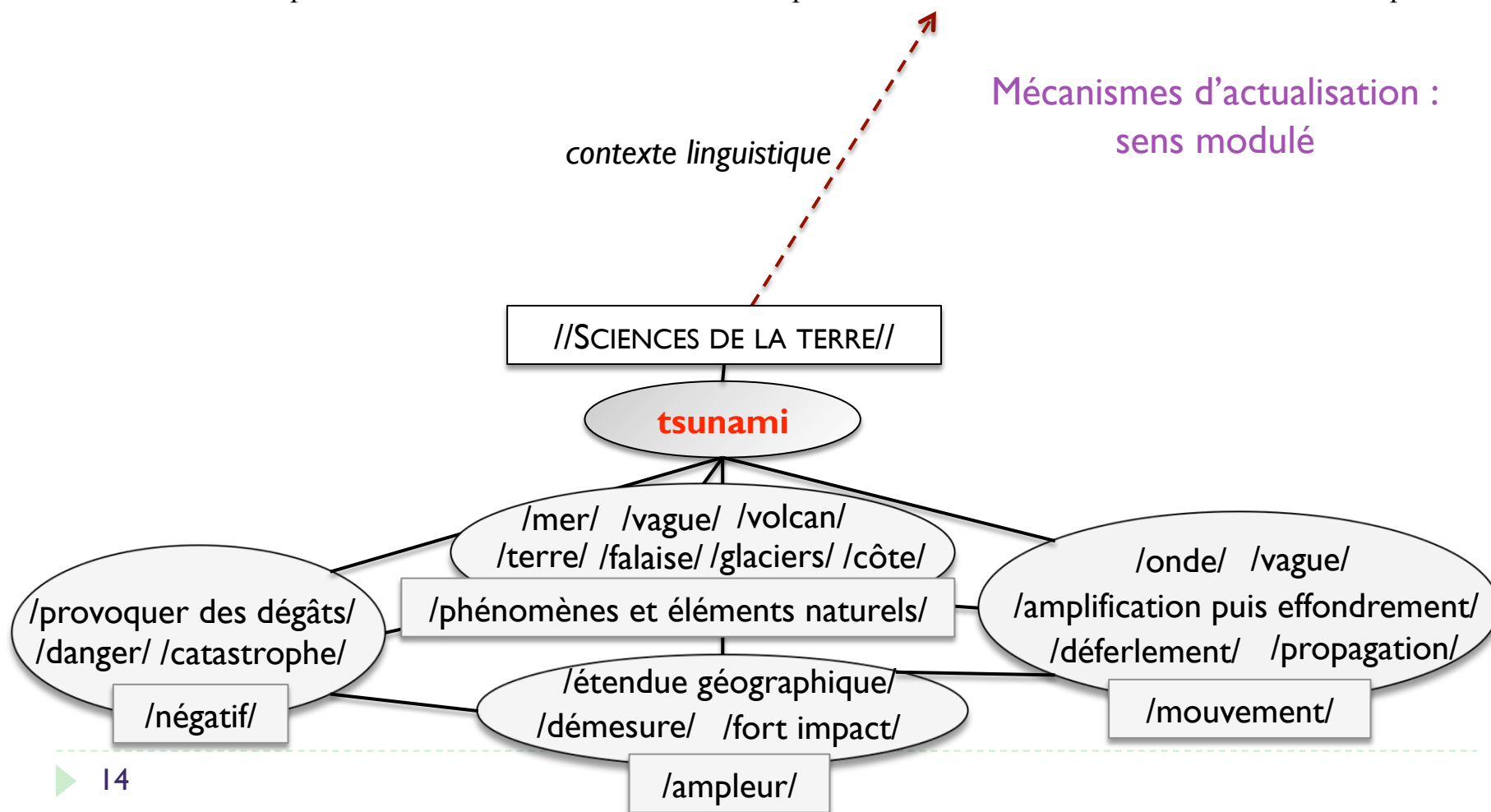
↳ *Représentation du sens :*
ensemble structuré de sèmes (traits sémantiques) = molécule sémique



Représentation du sens comme ensemble de sèmes (Rastier *et al.*, 1994) – (Rastier et Valette, 2009)

l'économie européenne vit dans la crainte d'être touchée par le **tsunami** financier arrivant d'outre-Atlantique

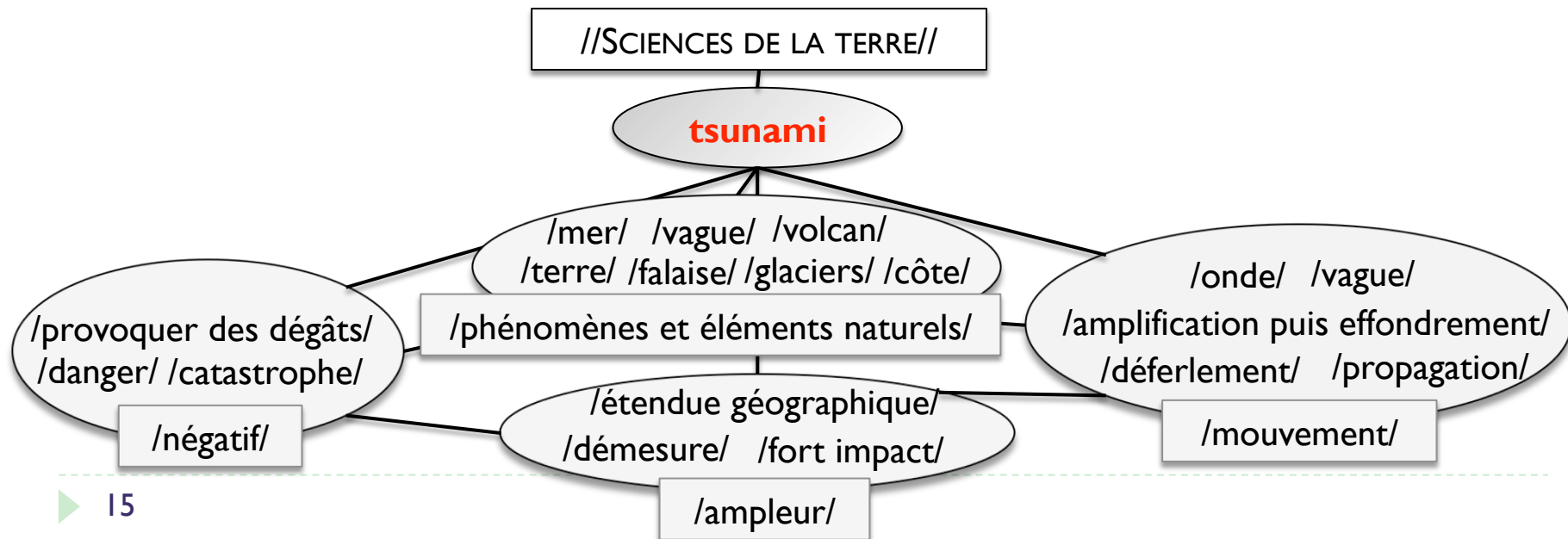
Mécanismes d'actualisation :
sens modulé



Mécanismes d'actualisation du sens en contexte

(Rastier, 2006) – (Valette, 2010)

Aujourd'hui, les excès du capitalisme casino l'ont presque conduit à sa ruine — menaçant l'économie réelle. En effet, l'économie européenne vit dans la crainte d'être touchée par le **tsunami** financier arrivant d'outre-Atlantique, en dépit de la relative solidité, jusqu'à présent, de la zone euro. Le gouvernement américain rassemble péniblement des centaines de milliards de dollars pour sauver les banques de leur propre folie ; la crise du crédit étrangle le financement du secteur industriel car les banques stockent l'argent pour se protéger. La récession menace.

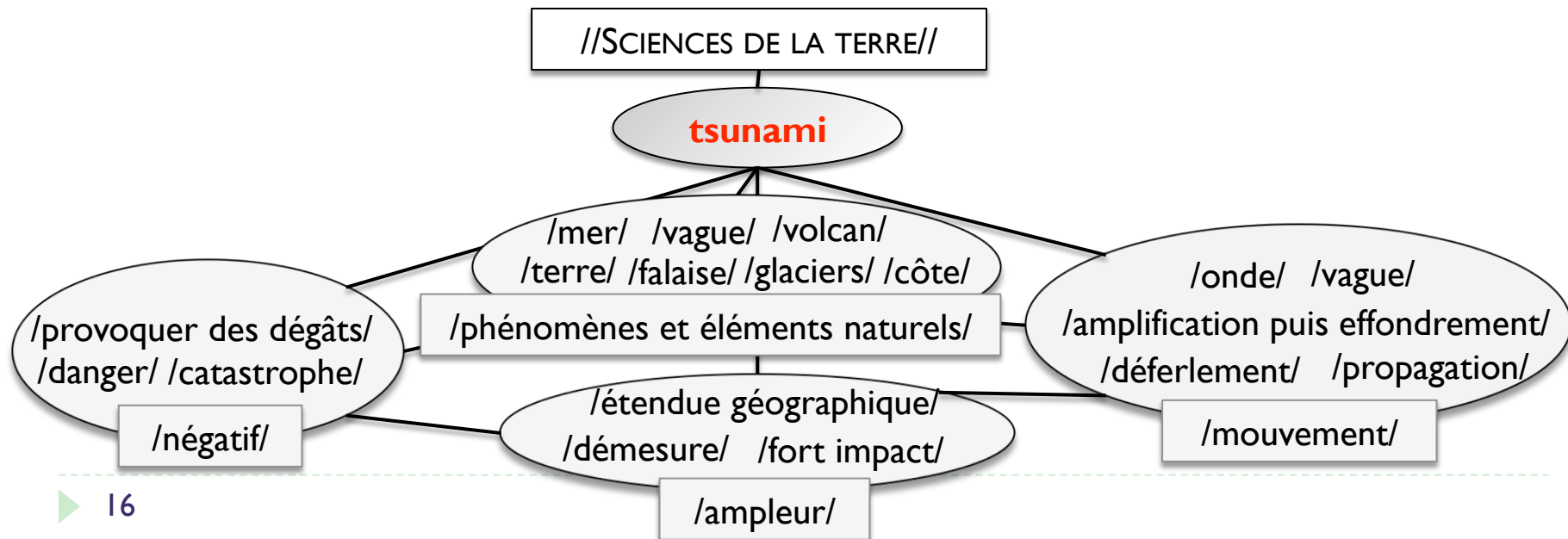


Mécanismes d'actualisation du sens : activation (Rastier, 2006) – (Valette, 2010)

Aujourd'hui, les excès du capitalisme casino l'ont presque conduit à sa **ruine** — **menaçant** l'économie réelle. En effet, /catastrophe/ /danger/

l'économie européenne vit dans la **crainte** d'être touchée par le **tsunami** financier arrivant d'outre-Atlantique, en /danger/ /provoquer des dégâts/

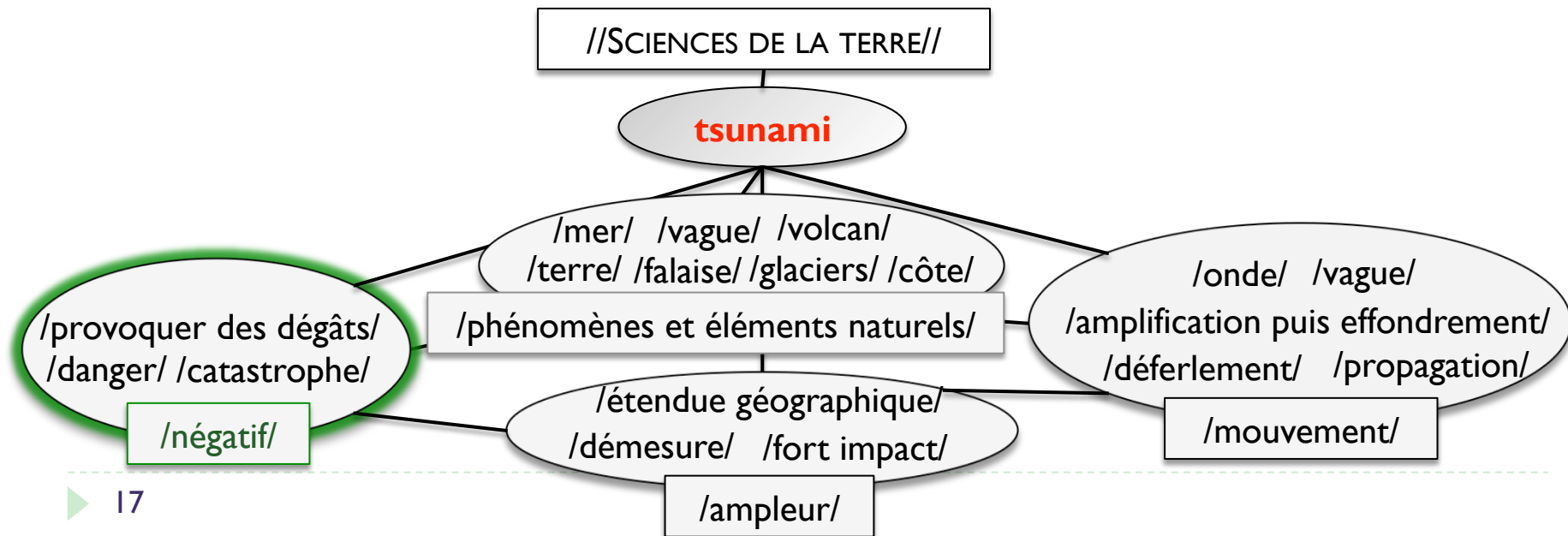
dépit de la relative solidité, jusqu'à présent, de la zone euro. Le gouvernement américain rassemble **péniblement** des centaines de milliards de dollars pour sauver les banques de leur propre folie ; la **crise** du crédit **étrangle** le /catastrophe//provoquer des dégâts/ financement du secteur industriel car les banques stockent l'argent pour **se protéger**. La récession **menace**. /danger/ /danger/



Mécanismes d'actualisation du sens : activation

(Rastier, 2006) – (Valette, 2010)

Aujourd'hui, les excès du capitalisme casino l'ont presque conduit à sa ruine — menaçant l'économie réelle. En effet, l'économie européenne vit dans la crainte d'être touchée par le tsunami financier arrivant d'outre-Atlantique, en dépit de la relative solidité, jusqu'à présent, de la zone euro. Le gouvernement américain rassemble péniblement des centaines de milliards de dollars pour sauver les banques de leur propre folie ; la crise du crédit étrangle le financement du secteur industriel car les banques stockent l'argent pour se protéger. La récession menace.



Mécanismes d'actualisation du sens : inhibition

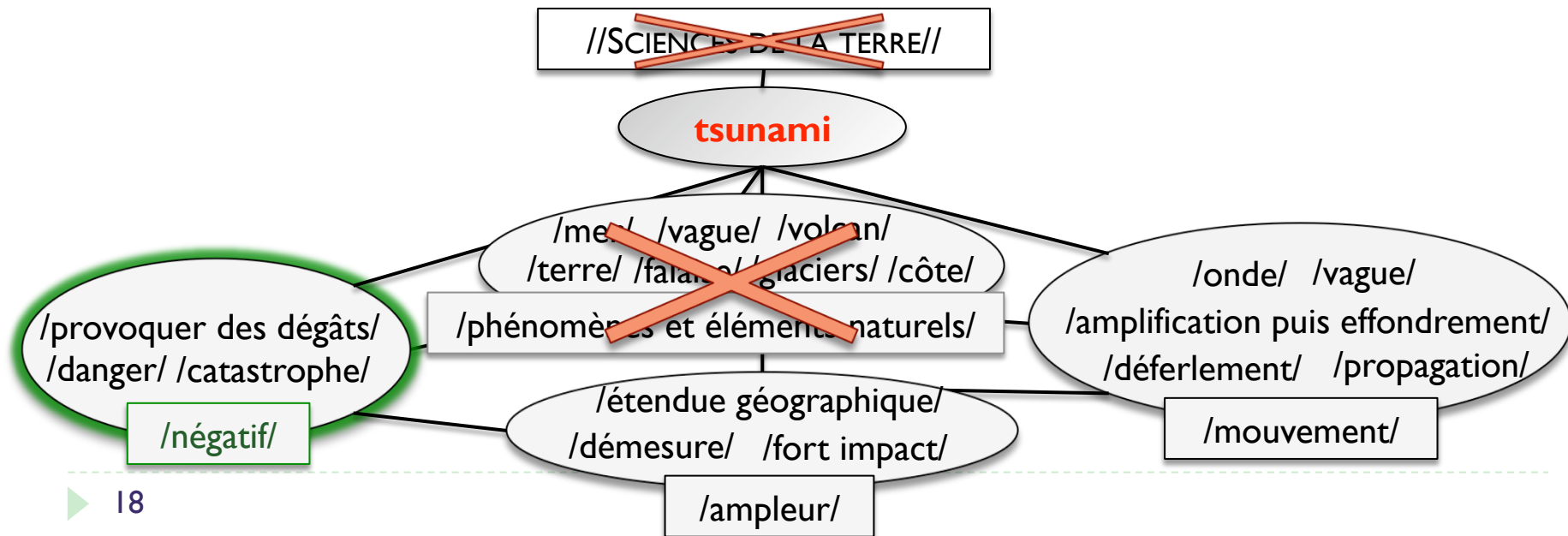
(Rastier, 2006) – (Valette, 2010)

Aujourd'hui, les excès du capitalisme casino l'ont presque conduit à sa **ruine** — **menaçant** l'économie réelle. En effet, /catastrophe/ /danger/

l'économie européenne vit dans la **crainte** d'être touchée par le **tsunami** financier arrivant d'outre-Atlantique, en /danger/ /provoquer des dégâts/

dépit de la relative solidité, jusqu'à présent, de la zone euro. Le gouvernement américain rassemble **péniblement** des centaines de milliards de dollars pour sauver les banques de leur propre folie ; la crise du crédit **étrangle** le /catastrophe//provoquer des dégâts/

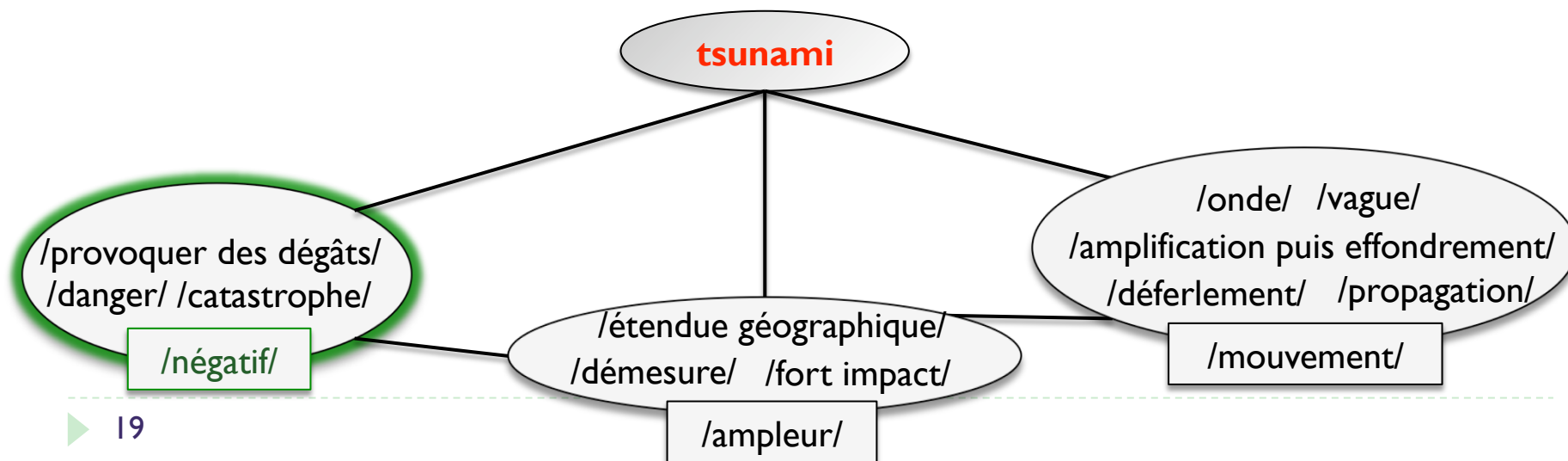
financement du secteur industriel car les banques stockent l'argent pour **se protéger**. La récession **menace**. /danger/ /danger/



Mécanismes d'actualisation du sens: enrichissement

(Rastier, 2006) – (Valette, 2010)

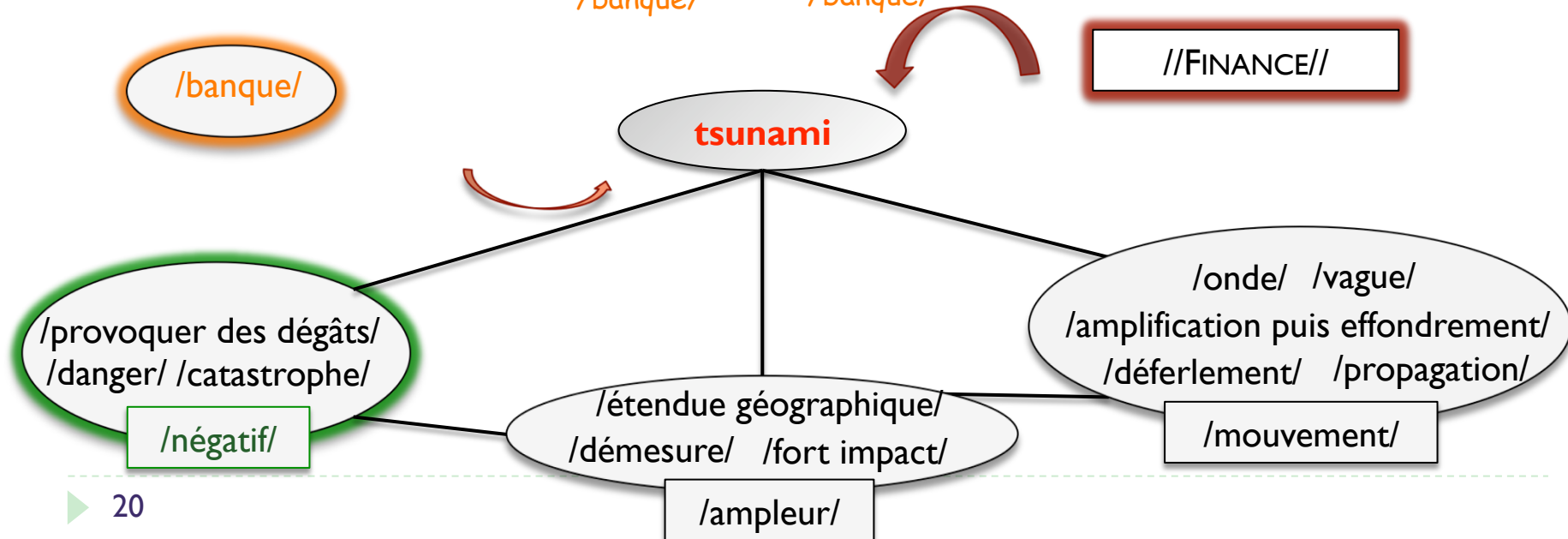
Aujourd'hui, les excès du **capitalisme** casino l'ont presque conduit à sa ruine — menaçant l'**économie réelle**. En effet, l'économie européenne vit dans la crainte d'être touchée par le **tsunami financier** arrivant d'outre-Atlantique, en dépit de la relative solidité, jusqu'à présent, de la **zone euro**. Le gouvernement américain rassemble péniblement des centaines de milliards de **dollars** pour sauver les **banques** de leur propre folie ; la crise du **crédit** étrangle le **financement** du secteur industriel car les **banques** stockent l'**argent** pour se protéger. La récession menace.



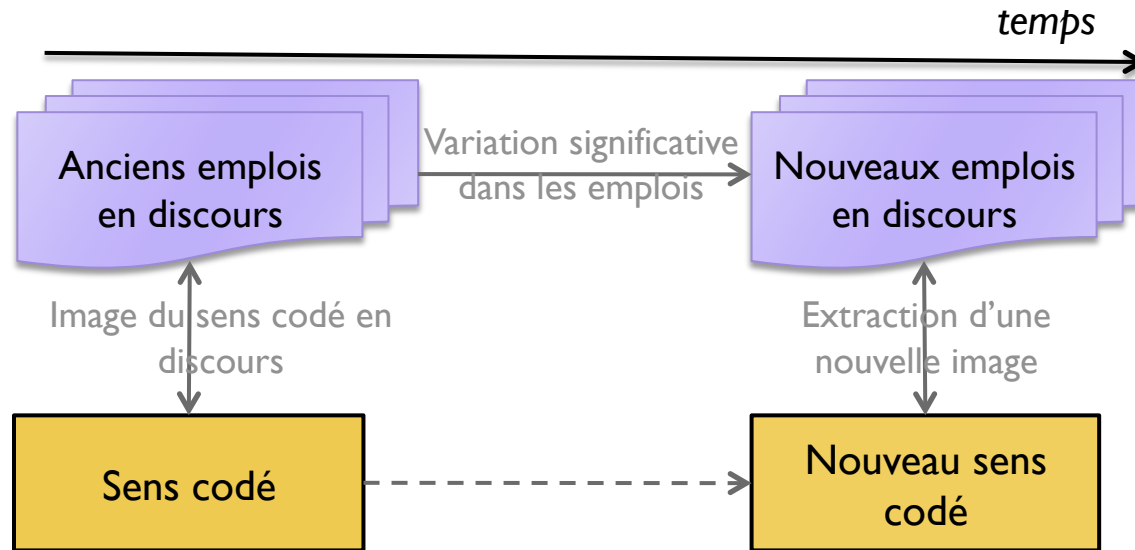
Mécanismes d'actualisation du sens: enrichissement

(Rastier, 2006) – (Valette, 2010)

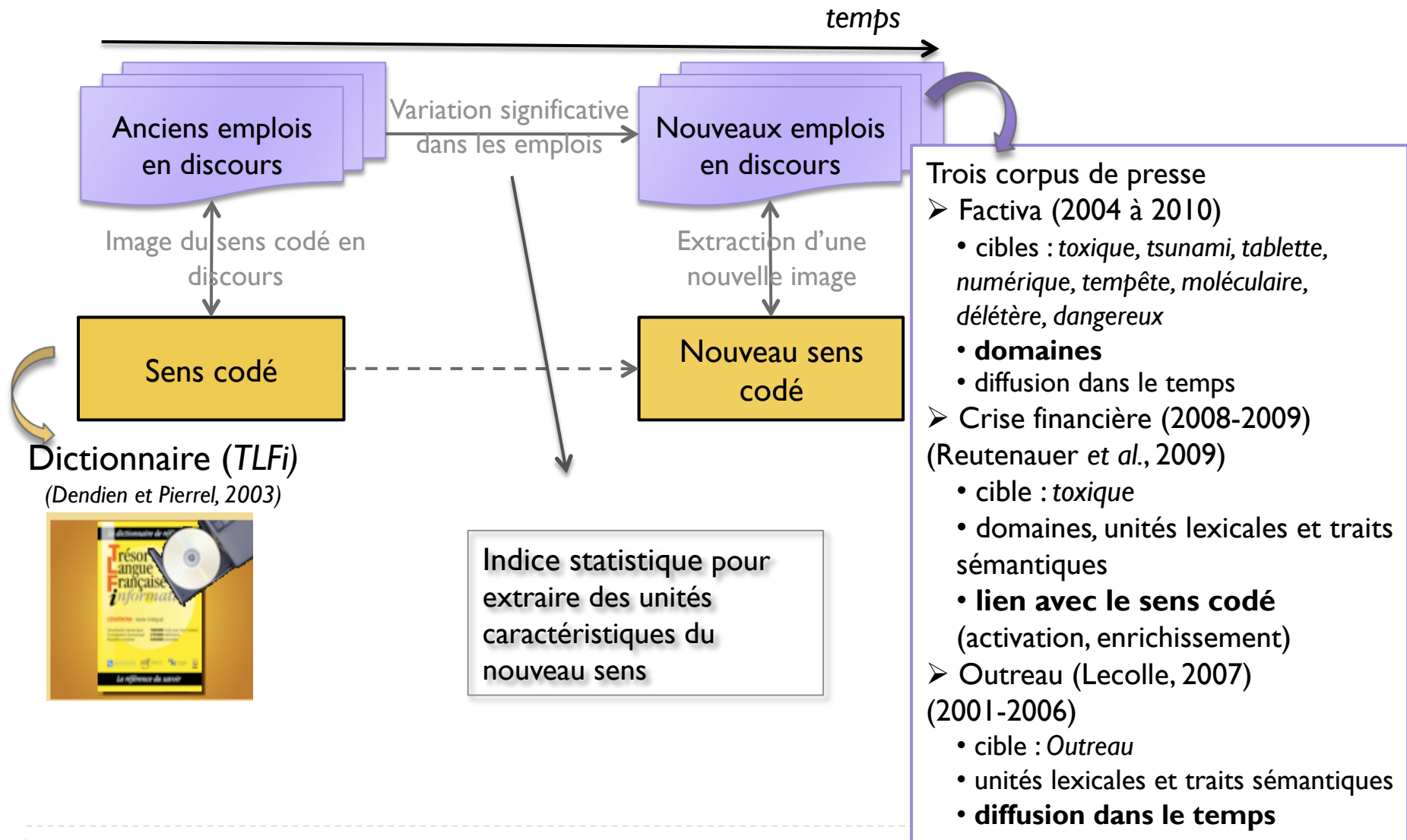
Aujourd'hui, les excès du **capitalisme** casino l'ont presque conduit à sa ruine — menaçant l'**économie réelle**. En effet, l'économie européenne vit dans la crainte d'être touchée par le **tsunami financier** arrivant d'outre-Atlantique, en dépit de la relative solidité, jusqu'à présent, de la **zone euro**. Le gouvernement américain rassemble péniblement des centaines de milliards de **dollars** pour sauver les **banques** de leur propre folie ; la crise du **crédit** étrangle le **financement** du secteur industriel car les **banques** stockent l'**argent** pour se protéger. La récession menace.



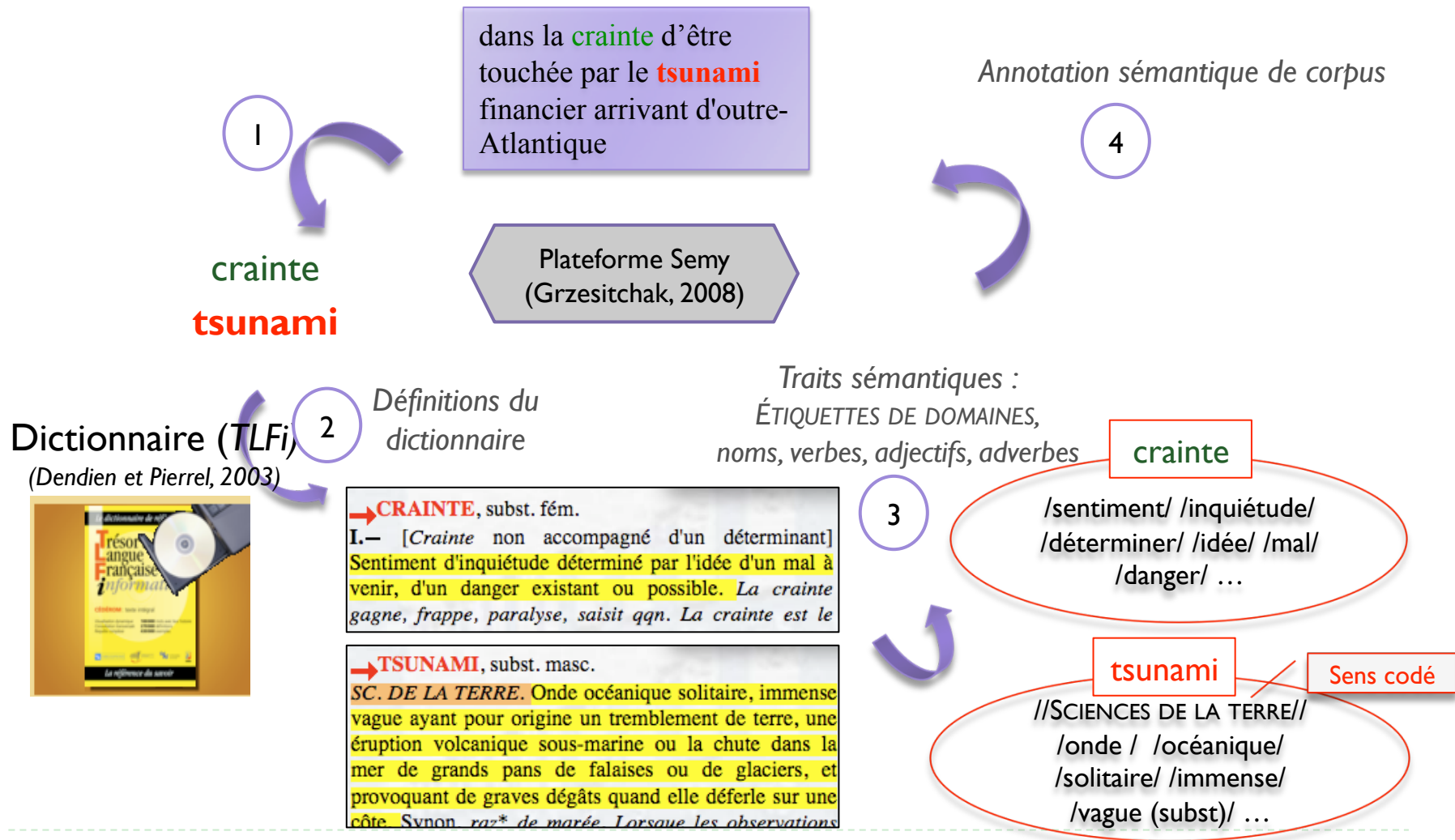
Vers un traitement automatique : structure de base d'un protocole



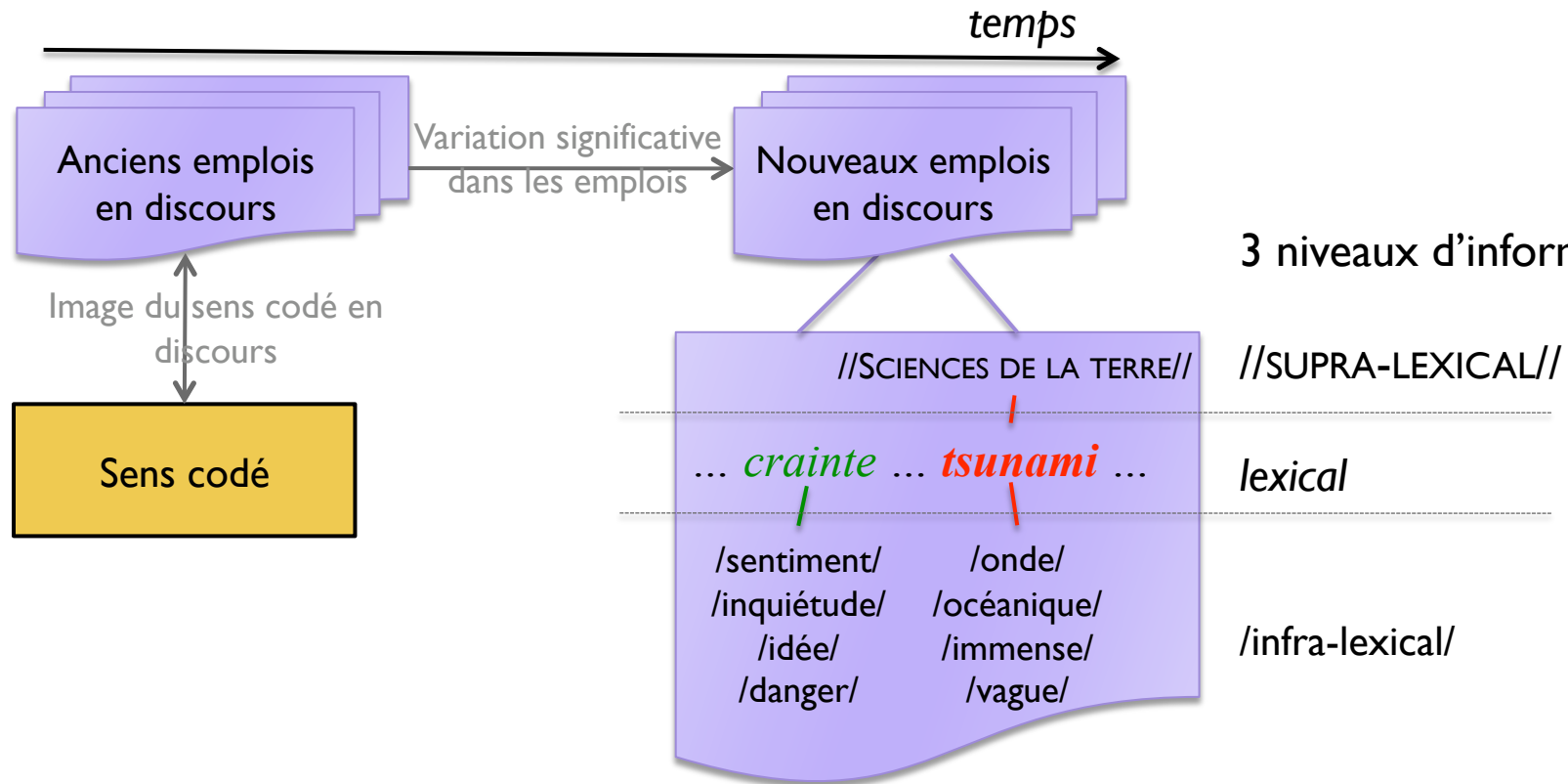
Vers un traitement automatique : structure de base d'un protocole



Exploitation du dictionnaire pour obtenir le sens codé et les unités infra- et supra-lexicales (Valette, 2008)



Trois niveaux d'observation après annotation



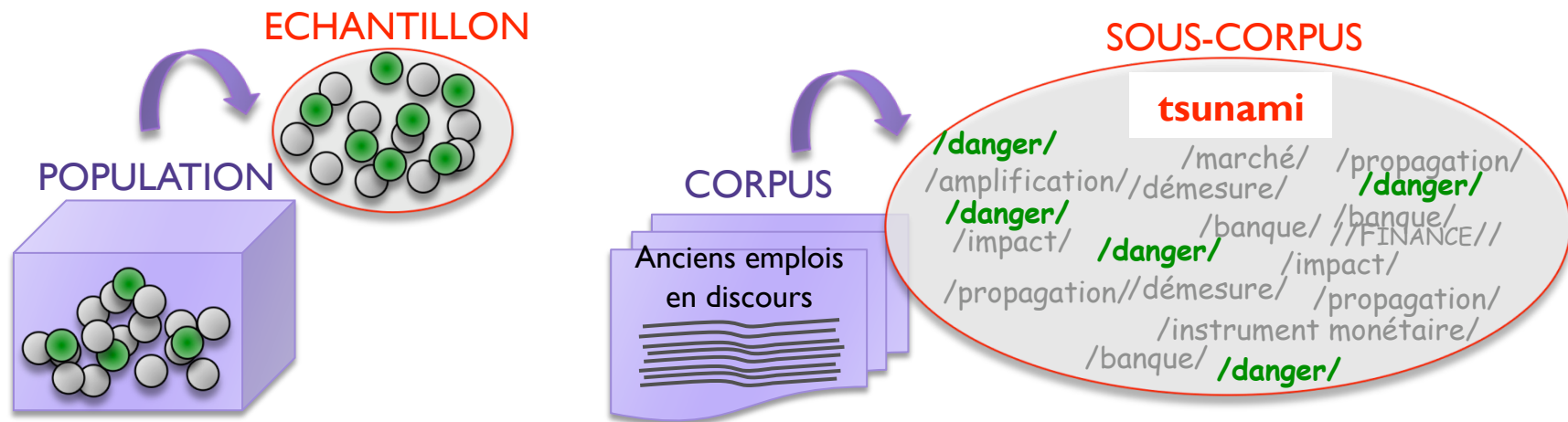
Unités caractéristiques du nouveau sens ?

Vers un traitement automatique : critère statistique pour extraire des unités

Hypothèse de distributionnalité (Harris, 1968), (Firth, 1957)

Des unités qui apparaissent dans les mêmes contextes sont de sens proche

Calcul d'un **degré d'affinité sémantique** avec la cible lexicale pour chaque unité (domaine, unité lexicale, trait sémantique)



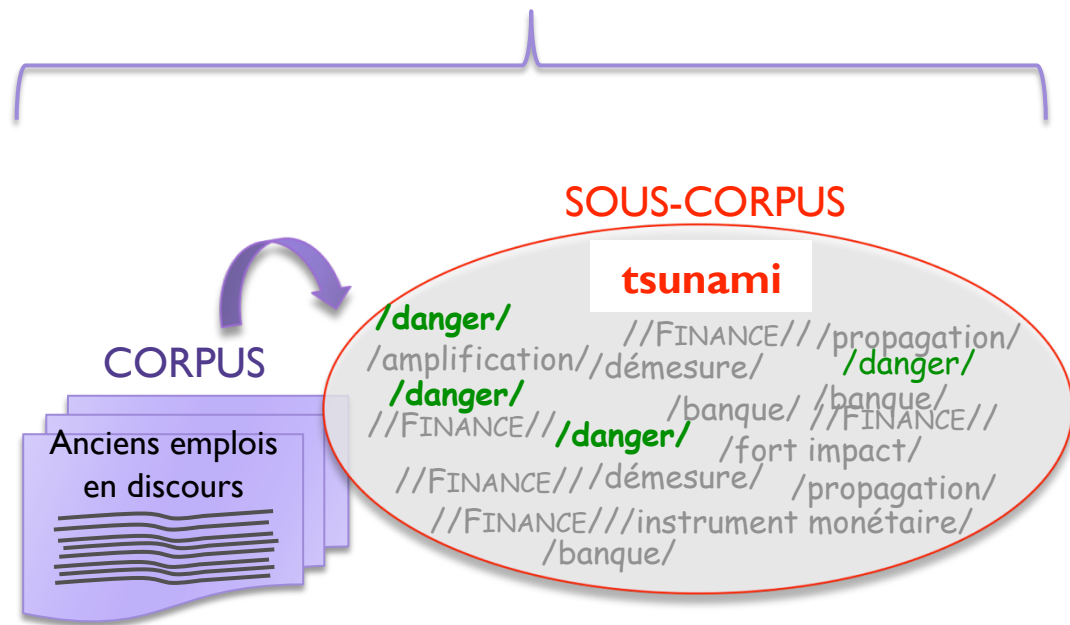
↳ Probabilité de /danger/ dans le sous-corpus

Indice de spécificité (Lafon, 1984), logiciel Lexico3 (Salem et al., 2003)

Vers un traitement automatique : critère statistique pour extraire des unités

SOUS-CORPUS / CORPUS

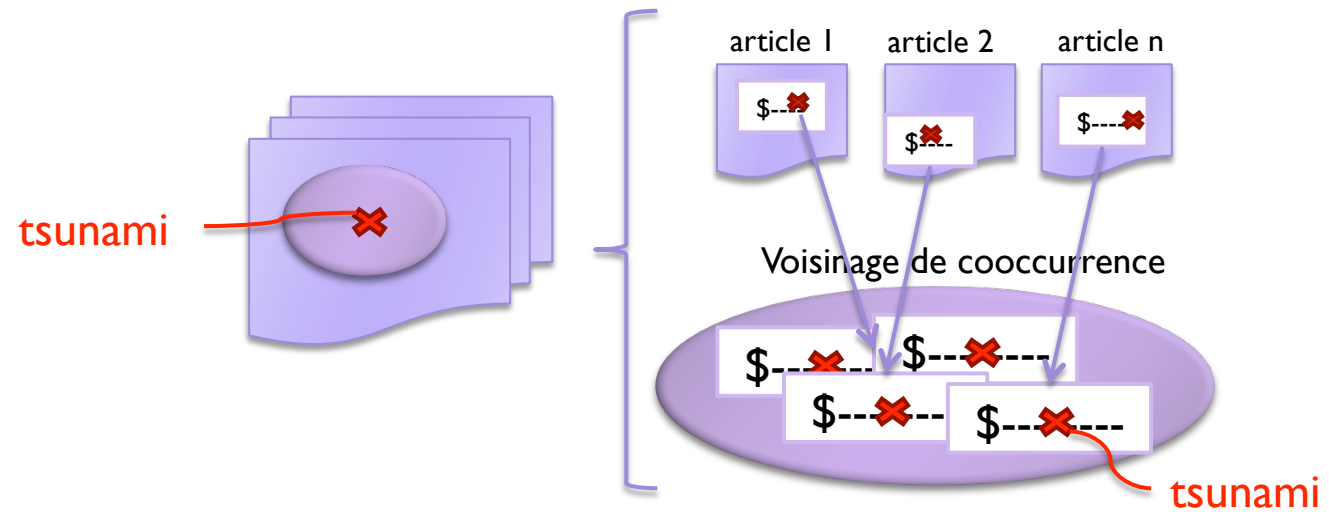
↳ quel découpage ?



Vers un traitement automatique : un espace textuel structuré (Mellet et Barthélémy, 2009)

Structuration de l'espace textuel en SOUS-CORPUS / CORPUS

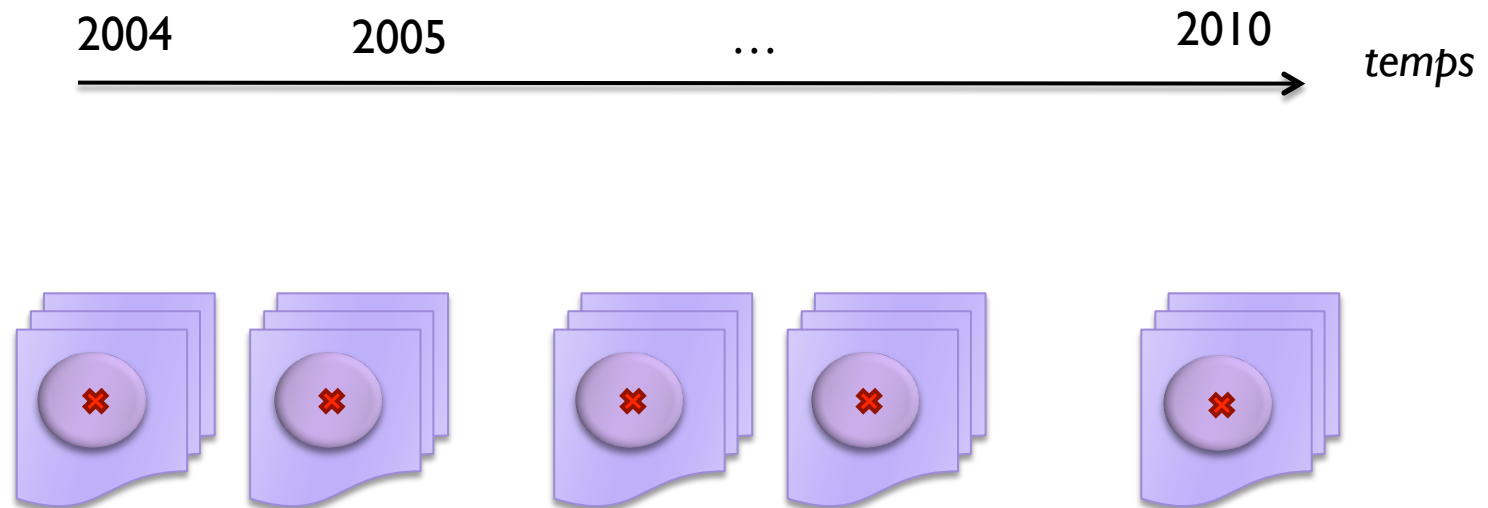
I. Voisinage de la cible



Vers un traitement automatique : un espace textuel structuré

Structuration de l'espace textuel en SOUS-CORPUS / CORPUS

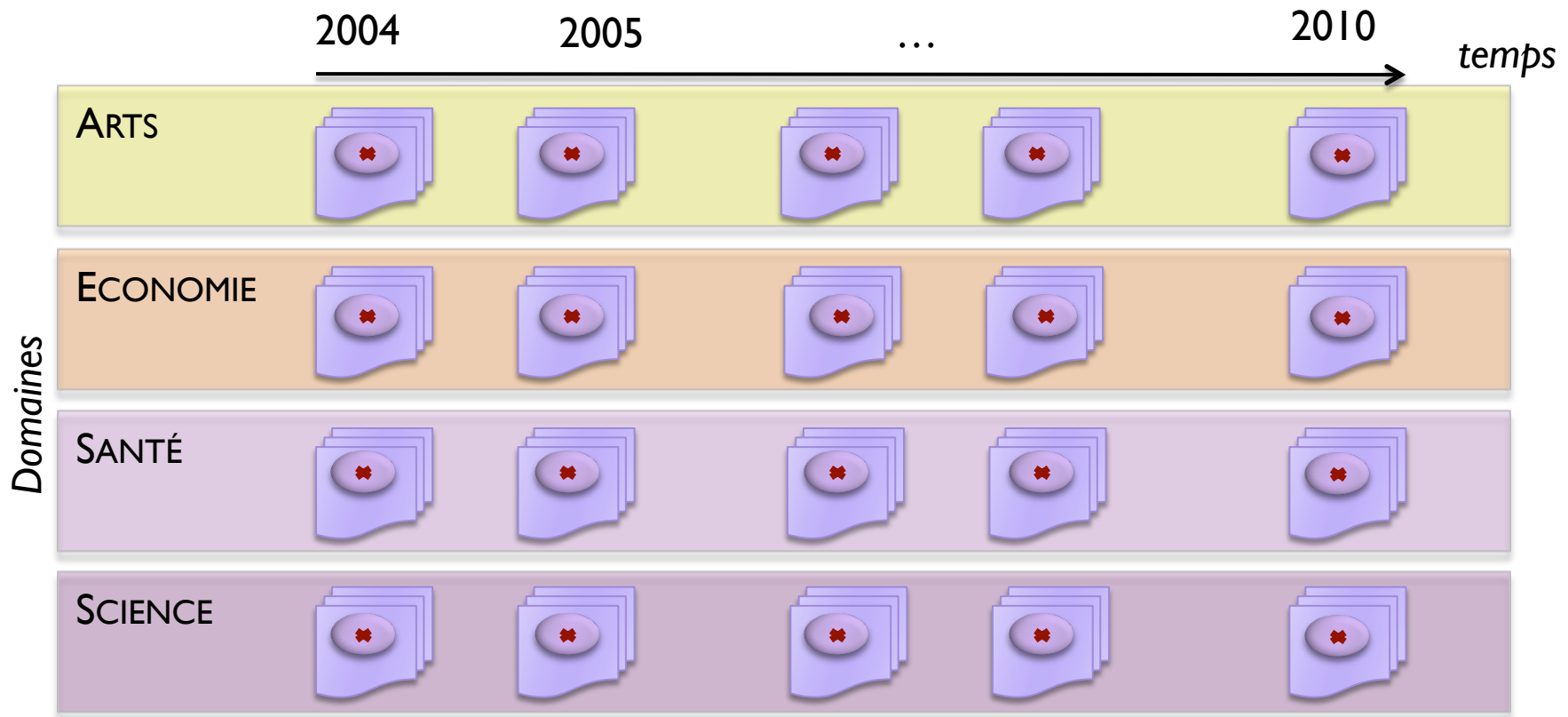
1. Voisinage de la cible
2. Temps



Vers un traitement automatique : un espace textuel structuré

Structuration de l'espace textuel en **SOUS-CORPUS** / **CORPUS**

1. Voisinage de la cible
2. Temps
3. Domaines



Vers un traitement automatique : une succession de jeux de contrastes

► Une procédure sous-tendue par des jeux de contrastes

I. Granularité sémantique

- a) DOMAINES
- b) *Unités lexicales*
- c) /Traits sémantiques/

2. Lien avec le sens codé

Unités appartenant ou non au sens codé (reconfiguration ou enrichissement du sens)

3. Paramètres textuels

- Temps : 2 périodes, plusieurs périodes
- Domaines
- Voisinages



Vers un traitement automatique : une succession de jeux de contrastes

► Une procédure sous-tendue par des jeux de contrastes

I. Granularité sémantique

a) **DOMAINES**

b) *Unités lexicales*

c) */Traits sémantiques/*

2. Lien avec le sens codé

Unités appartenant ou non au sens codé (reconfiguration ou enrichissement du sens)

3. Paramètres textuels

- Temps : 2 périodes, plusieurs périodes
- Domaines
- Voisinages

Diffusion en corpus

toxique

tablette

numérique

dangereux délétère

tsunami

moléculaire

tempête

raz-de-marée

Vers un traitement automatique : une succession de jeux de contrastes

► Une procédure sous-tendue par des jeux de contrastes

I. Granularité sémantique

- a) **DOMAINES**
- b) *Unités lexicales*
- c) */Traits sémantiques/*

2. Lien avec le sens codé

Unités appartenant ou non au sens codé (reconfiguration ou enrichissement du sens)

3. Paramètres textuels

- Temps : **2 périodes**, plusieurs périodes
- Domaines
- Voisinages

Diffusion en corpus

toxique *tablette*
tsunami *moléculaire*

numérique
tempête

dangereux *délétère*
raz-de-marée

Tri des cibles

toxique
tsunami
tablette
~~*moléculaire*~~
numérique
tempête
~~*dangereux*~~
~~*délétère*~~
~~*raz-de-marée*~~

Affectation de domaines

toxique
ÉCONOMIE
numérique
domaines multiples

Vers un traitement automatique : une succession de jeux de contrastes

► Une procédure sous-tendue par des jeux de contrastes

I. Granularité sémantique

- a) **DOMAINES**
- b) *Unités lexicales*
- c) */Traits sémantiques/*

2. Lien avec le sens codé

Unités appartenant ou non au sens codé (reconfiguration ou enrichissement du sens)

3. Paramètres textuels

- Temps : 2 périodes, plusieurs périodes
- Domaines
- Voisinages

Diffusion en corpus

toxique *tablette*
tsunami *moléculaire*

numérique
tempête

dangereux *délétaire*
raz-de-marée

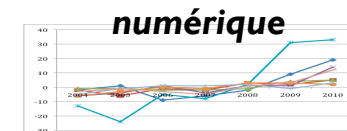
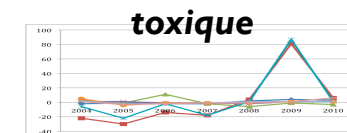
Tri des cibles

toxique
tsunami
tablette
~~*moléculaire*~~
~~*numérique*~~
~~*tempête*~~
~~*dangereux*~~
~~*délétaire*~~
~~*raz-de-marée*~~

Affectation de domaines

toxique
ÉCONOMIE
numérique
domaines multiples

Régularité d'évolution



Vers un traitement automatique : une succession de jeux de contrastes

► Une procédure sous-tendue par des jeux de contrastes

I. Granularité sémantique

- a) **DOMAINES**
- b) *Unités lexicales*
- c) /Traits sémantiques/

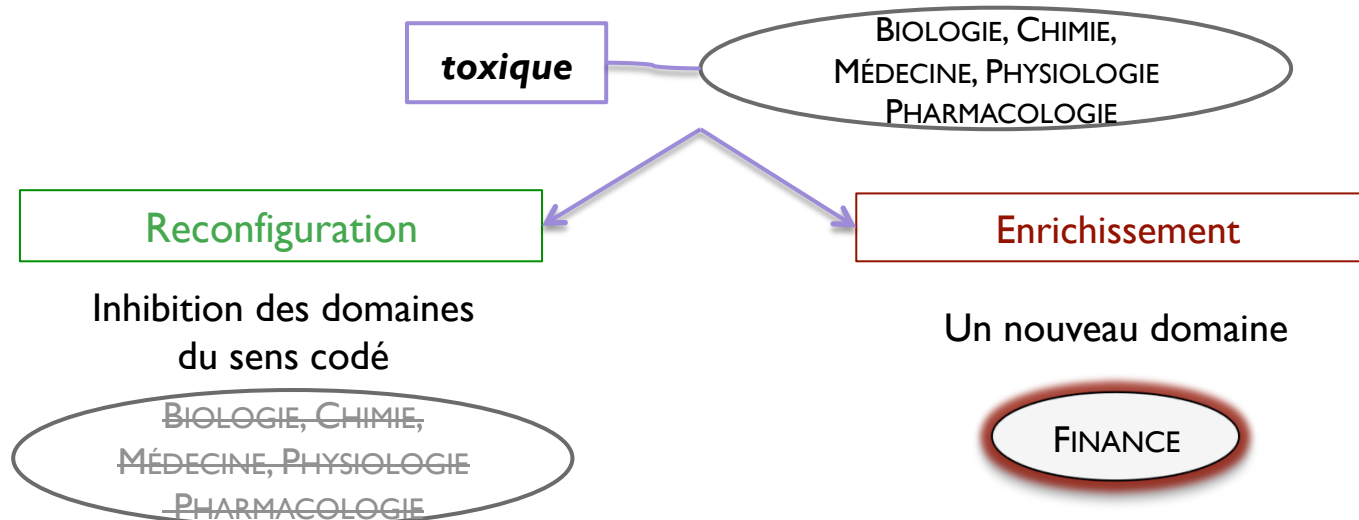
2. Lien avec le sens codé

Unités appartenant ou non au sens codé (**reconfiguration** ou **enrichissement** du sens)

3. Paramètres textuels

- Temps : 2 périodes, plusieurs périodes
- Domaines
- Voisinages

Lien avec le sens codé



Vers un traitement automatique : une succession de jeux de contrastes

► Une procédure sous-tendue par des jeux de contrastes

I. Granularité sémantique

- a) DOMAINES
- b) *Unités lexicales*
- c) **/Traits sémantiques/**

2. Lien avec le sens codé

Unités appartenant ou non au sens codé (**reconfiguration** ou **enrichissement** du sens)

3. Paramètres textuels

- Temps : 2 périodes, plusieurs périodes
- Domaines
- Voisinages

Reconfiguration

toxique

/produit/, /vivre/,
/minéral/ /animal/, /végétal/
/provoquer/, /intoxication/,
/organisme/, /destruction/,...

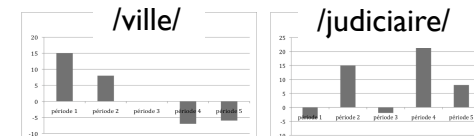
Inhibition des traits
du sens codé

/minéral/ -3
/vivre/ -4

Outreau

/ville/, /Pas-de-Calais/, /français/
/croyance/, /existence/, /réseau/,
/pédophile/, /erreur/, /judiciaire/
/réfutation/, ...

Modulation dans le temps :
convergence avec l'analyse
de Lecolle (2007)



Vers un traitement automatique : une succession de jeux de contrastes

► Une procédure sous-tendue par des jeux de contrastes

I. Granularité sémantique

- a) DOMAINES
- b) *Unités lexicales*
- c) **/Traits sémantiques/**

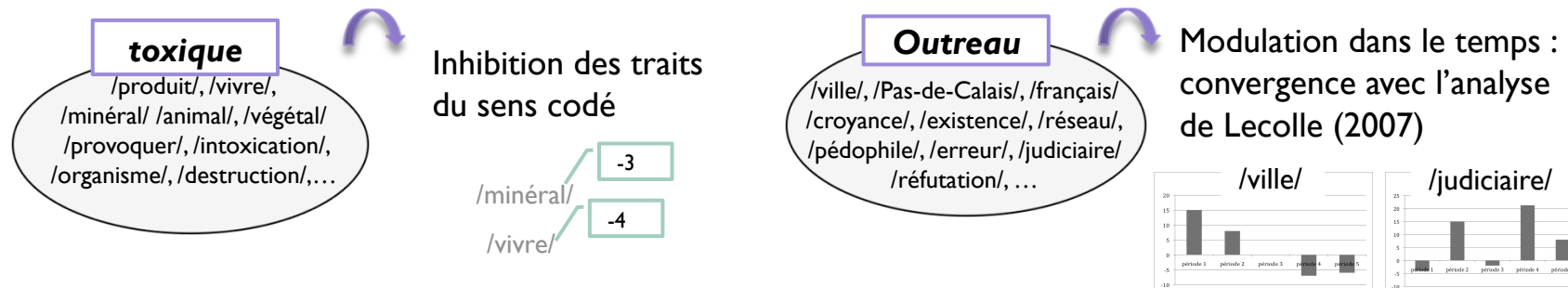
2. Lien avec le sens codé

Unités appartenant ou non au sens codé (**reconfiguration** ou **enrichissement** du sens)

3. Paramètres textuels

- Temps : 2 périodes, plusieurs périodes
- Domaines
- Voisinages

Reconfiguration



Enrichissement

toxique

Critères : traits les plus spécifiques et dépendants des domaines émergents

↳ Traits saillants dépendants de FINANCE : /banque/, /bilan/, /prêt/, ...

Bilan

Quel phénomène linguistique ?

Nouveauté par rapport à quoi ?

Il est possible de construire de façon semi-automatique une représentation structurée et nuancée d'un nouveau sens lexical

Où chercher l'information ?

Quel modèle de représentation ?

Comment ?

Bilan

Quel phénomène linguistique ?

Néosémie

Nouveauté par rapport à quoi ?

- Temps
- Sens codé vs emplois en discours

Il est possible de construire de façon semi-automatique une représentation structurée et nuancée d'un nouveau sens lexical

Où chercher l'information ?

Quel modèle de représentation ?

Comment ?

Bilan

Quel phénomène linguistique ?

Néosémie

Nouveauté par rapport à quoi ?

- Temps
- Sens codé vs emplois en discours

Il est possible de construire de façon semi-automatique une représentation structurée et nuancée d'un nouveau sens lexical

Où chercher l'information ?

- Global et local
- Palier de textualité large
- Cooccurents

Quel modèle de représentation ?

- Inspiré de la sémantique textuelle
- Niveaux de description :
infra-lexical / lexical / supra-lexical
 - Traits sémantiques
 - ✓ Réccurrence de traits
 - ✓ Activation / inhibition / enrichissement

Comment ?

Protocole :

- Outils statistiques
- Information sémantique affinée successivement (granularité sémantique décroissante)
- Jeu sur l'espace textuel
- Interactions entre sens codé et sens en discours

Bilan

Quel phénomène linguistique ?

Néosémie

Nouveauté par rapport à quoi ?

- Temps
- Sens codé vs emplois en discours

Il est possible de construire de façon semi-automatique une représentation structurée et nuancée d'un nouveau sens lexical

Où chercher l'information ?

- **Global et local**
- Palier de textualité large
- **Cooccurents**

Quel modèle de représentation ?

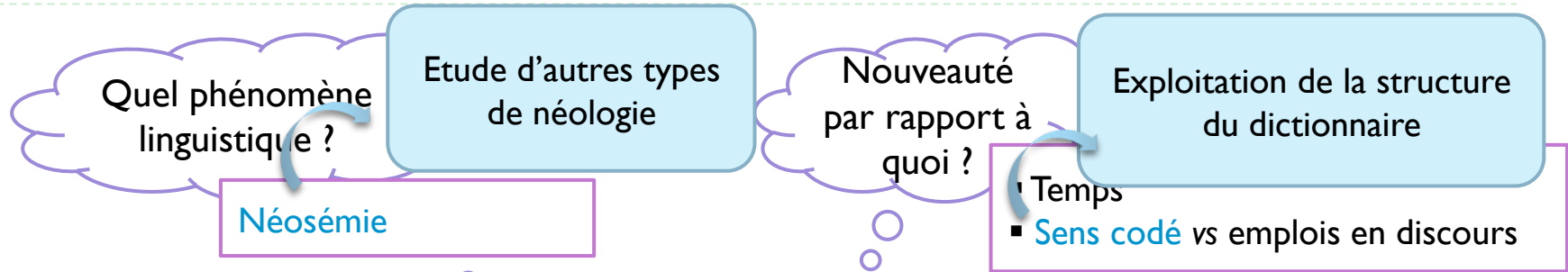
- Inspiré de la sémantique textuelle
- Niveaux de description :
infra-lexical / lexical / **supra-lexical**
- Traits sémantiques
 - ✓ Récurrence de traits
 - ✓ Activation / inhibition / enrichissement

Comment ?

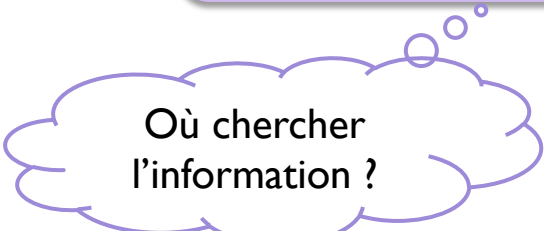
Protocole :

- Outils **statistiques**
- **Information sémantique** affinée successivement (**granularité sémantique décroissante**)
- Jeu sur l'espace textuel
- Interactions entre sens codé et sens en discours

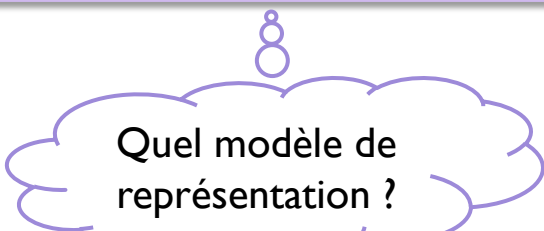
Bilan



Il est possible de construire de façon semi-automatique une représentation structurée et nuancée d'un nouveau sens lexical



- Global et local
- Palier de textualité large
- Cooccurents



- Inspiré de la sémantique textuelle
- Niveaux de description :
infra-lexical / lexical / supra-lexical
- Traits sémantiques
 - ✓ Récurrence de traits
 - ✓ Activation / inhibition / enrichissement



- Protocole :
- Outils statistiques
 - Information sémantique affinée successivement (granularité sémantique décroissante)
 - Jeu sur l'espace textuel
 - Interactions entre sens codé et sens en discours

Perspectives

Etude d'autres types de néologie

hameçonnage

Analyse morphologique

Action de + hameçonner

Résultat de l'action de

Procédure utilisée pour la néosémie

Sens codé



/action/
+ traits de
hameçonner

/ action/ /garnir/ /hameçon/
/attirer/ /apparence/
/trompeur/

Test en corpus

Emplois de
hameçonnage

//INFORMATIQUE//
/web/ /courriel/

réalité augmentée

Analyse du figement

Pré-construction
d'un sens codé

réalité + augmenter



sens codé de réalité
+ sens codé d'augmenter

Emplois de
réalité augmentée

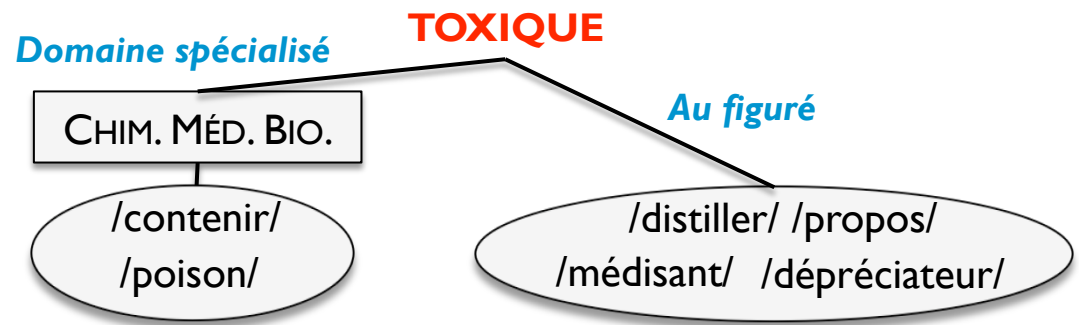
Sens codé retravaillé :
- traits éliminés

- ajout de traits absents des deux entrées

Perspectives

Exploitation de la structure du dictionnaire

TOXIQUE
 ➤ CHIM. MED. BIOL.. Qui est ou qui contient un poison
 ➤ *Au fig.* – Qui distille plus ou moins ouvertement des propos médisants ou dépréciateurs

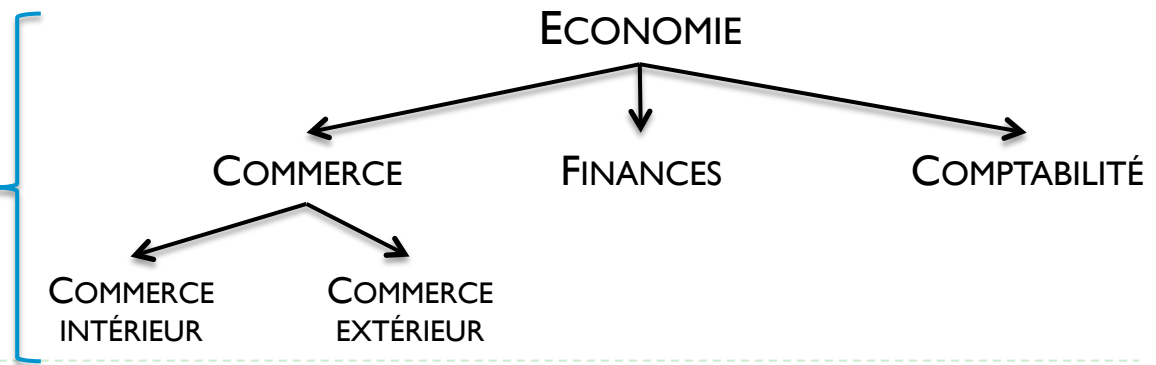


➔ 2 sous-ensembles au lieu d'un seul sac de traits
 ➔ poids différents ?



Hierarchie à plusieurs niveaux

➔ affinements plus progressifs
 ➔ regroupements d'entrées



Bibliographie

- ▶ BASTUJI J. (1974). « Aspects de la néologie sémantique », *Langages*, n°36.
- ▶ BAUER L., RENOUF A. (2000). « Contextual clues to Word-Meaning », *International Journal of corpus Linguistics*, Vol. 5, n°2, 231-259.
- ▶ DURY P. (2008). « Les noms du pétrole : une approche diachronique de la métonymie onomastique », *Lexis, E-Journal in English Lexicology*, <http://screcherche.univ-lyon3.fr/lexis/>.
- ▶ FIRTH J. R. (1957). *Papers in Linguistics 1934-1951*. London: Oxford University Press.
- ▶ GUILBERT, L. (1974). « Grammaire générative et néologie lexicale », *Langages*, vol. 8, n°36.
- ▶ HARRIS Z. S. (1968). *Mathematical Structures of Languages*. New York : John Wiley & Sons.
- ▶ LAFON P. (1984). *Dépouillements et statistiques en lexicométrie*. Genève-Paris, éd. Slatkine – Champion, pp.54-77.
- ▶ LECOLLE M. (2007). « Polysignifiante du toponyme, historicité du sens et interprétation en corpus », *Corpus*, n°6, 101-125.
- ▶ MEJRI Salah (2010). « Néologie et traitement automatique », in *Actes del i Congrès internacional de neologia de les llengües romàniques (CINEO 2008)*, Barcelone, Espagne, 7-10 mai 2008, pp.99-110.
- ▶ MELLET S., BARTHÉLEMY J.P. (2009). « La topologie textuelle : légitimation d'une notion émergente », *Lexicometrica 7, Topographie et topologie textuelles*.
- ▶ PALLAUD B. (2004). « La transgression et la variation », *Marges linguistiques*, n°8 *Langues, langage, Inconscient, Linguistique et Psychanalyse*, vol. 2, 76-87.
- ▶ PICTON A. (2009). *Diachronie en langue de spécialité. Définition d'une méthode linguistique outillée pour repérer l'évolution des connaissances en corpus. Un exemple appliqué au domaine spatial*. Thèse de Doctorat en Sciences du Langage. Université Toulouse 2.
- ▶ PRUVOST J., SABLAYROLLES J.-F. (2003). *Les néologismes*, PUF, Coll. Que sais-je?, Paris, 2003.
- ▶ RASTIER François (2006). « Formes sémantiques et textualité », *Langages*, 2006/3 n°163, (éd) Armand Colin, 99-114.
- ▶ RASTIER F., VALETTE M. (2009). « De la polysémie à la néosémie », *Le français moderne*, S. Mejri (éd.), *La problématique du mot*, 77, 97-116.
- ▶ ROUSSEAU J.-L., DEPECKER L. (1999). « Nouveaux outils pour la néologie », *Terminologies nouvelles*, n°20,, 2-3.
- ▶ SABLAYROLLES Jean-François (2000). *La néologie en français contemporain : examen du concept et analyse de productions néologiques récentes*. Champion, Paris.
- ▶ SALEM A., LAMALLE C., MARTINEZ W., FLEURY S., FRACCHIOLLA B., KUNCOVA A., MAISONDIEU A. (2003). « Lexico3 – Outils de statistique textuelle. Manuel d'utilisation. », Syled-CLA2T, Université de la Sorbonne nouvelle – Paris 3 : <http://www.cavi.univ-paris3.fr/llpga/ilpga/tal/lexicoWWW>.
- ▶ VALETTE M. (2008). « À quoi servent les lexiques sémantiques ? Discussion et proposition », *Description linguistique pour le traitement automatique du français, Cahiers du CENTAL*, n°5 – décembre 2008, Presses Universitaires de Louvain, 43-58.
- ▶ VALETTE M. (2010) « Méthodes pour la veille lexicale », in : *Actes de la journée d'étude Le dictionnaire électronique. Quelles perspectives pour les sciences humaines et sociales ?*, Leila Messaoudi (éds.), Publication du laboratoire Langage et société, Université Ibn Tofail Kénitra.