



Extraction semi-supervisée de couples d'antonymes grâce à leur morphologie

Didier Schwab, Mathieu Lafourcade et Violaine Prince.



Lirmm-info
Montpellier-France
équipe TALN



Situation (1/2)

Cadre : représentation du sens en TALN

Développement d'un système permettant la représentation du sens

Création d'outils exploitant cette représentation

Applications : toutes celles où la représentation du sens peut jouer un rôle

RI, traduction, résumé automatique, ...



Situation (2/2)

Moyens : exploitation des relations sémantiques existant entre les items lexicaux

Modélisation des relations sémantiques :

Synonymie (TALN'2001)

Hyperonymie(JADT'2004)

Antonymie (TALN'2002)

Problème : manque de sources gratuites disponibles pour l'antonymie

Solution : essayer de les fabriquer nous-mêmes

Comment ? En utilisant les oppositions morphologiques entre les antonymes (anti|||ε, mono|||poly)



L'antonymie

Deux items lexicaux sont en relation d'antonymie si on peut exhiber une symétrie de leurs traits sémantiques par rapport à un axe.

La symétrie peut se décliner en trois catégories suivant le support => trois types d'antonymie



Antonymie complémentaire

Symétrie de valeurs dans un système booléen

présence $\| \| \|_c$ absence, mort $\| \| \|_c$ vivant, pair $\| \| \|_c$ impair

Il est présent \Leftrightarrow il n'est pas absent

Il est absent \Leftrightarrow il n'est pas présent



Antonymie scalaire

Valeurs continues

Symétrie par rapport à une valeur de référence

chaud \parallel_s froid, grand \parallel_s petit

Cet homme est grand \Rightarrow Cet homme n'est pas petit

Cet homme est petit \Rightarrow Cet homme n'est pas grand

Cet homme n'est pas grand \Rightarrow Cet homme est petit ou de taille moyenne

Cet homme n'est pas petit \Rightarrow Cet homme est grand ou de taille moyenne



Antonymie duale

Duals conversifs

sémantique identique et inversion des actants syntaxiques

acheter $\| \| \|_d$ *vendre*, *prêter* $\| \| \|_d$ *emprunter*, *mari* $\| \| \|_d$ *femme*

Pierre est le père de Marc \Leftrightarrow *Marc est le fils de Pierre*

Duals propres

culturels ou spatio-temporels

question $\| \| \|_d$ *réponse*, *départ* $\| \| \|_d$ *arrivée*, *soleil* $\| \| \|_d$ *lune*



Morphologie et antonymie

Les préfixes sont parfois porteurs de sens

bi- signifie *deux* (*binaire, bidirectionnel*)

péri- correspond à une idée de tour (*périmètre*)

hémi- à une idée de moitié (*hémisphère, hémicycle*)

Les préfixes peuvent ainsi s'opposer

mono- (*un*) s'oppose à *poly-* (*plusieurs*)

hyper- (*au dessus*) /// *hypo-* (*au dessous*)

Recherche des antonymes basée sur ces oppositions
dans une liste d'items lexicaux



Méthode

Inspirée par les travaux d'Emmanuel Morin sur l'acquisition de schémas lexico-sémantiques

Deux différences :

- L'objectif premier est de construire une liste de couples d'antonymes (et **non** de chercher quels couples de préfixes s'opposent en langue).
- Les experts ne valident pas directement les schémas. Ils ne sont validés que si au moins UN couple d'items antonymes les valide (validation indirecte)



Processus d'acquisition

Fournir une liste
d'items antonymes

①

Processus d'acquisition

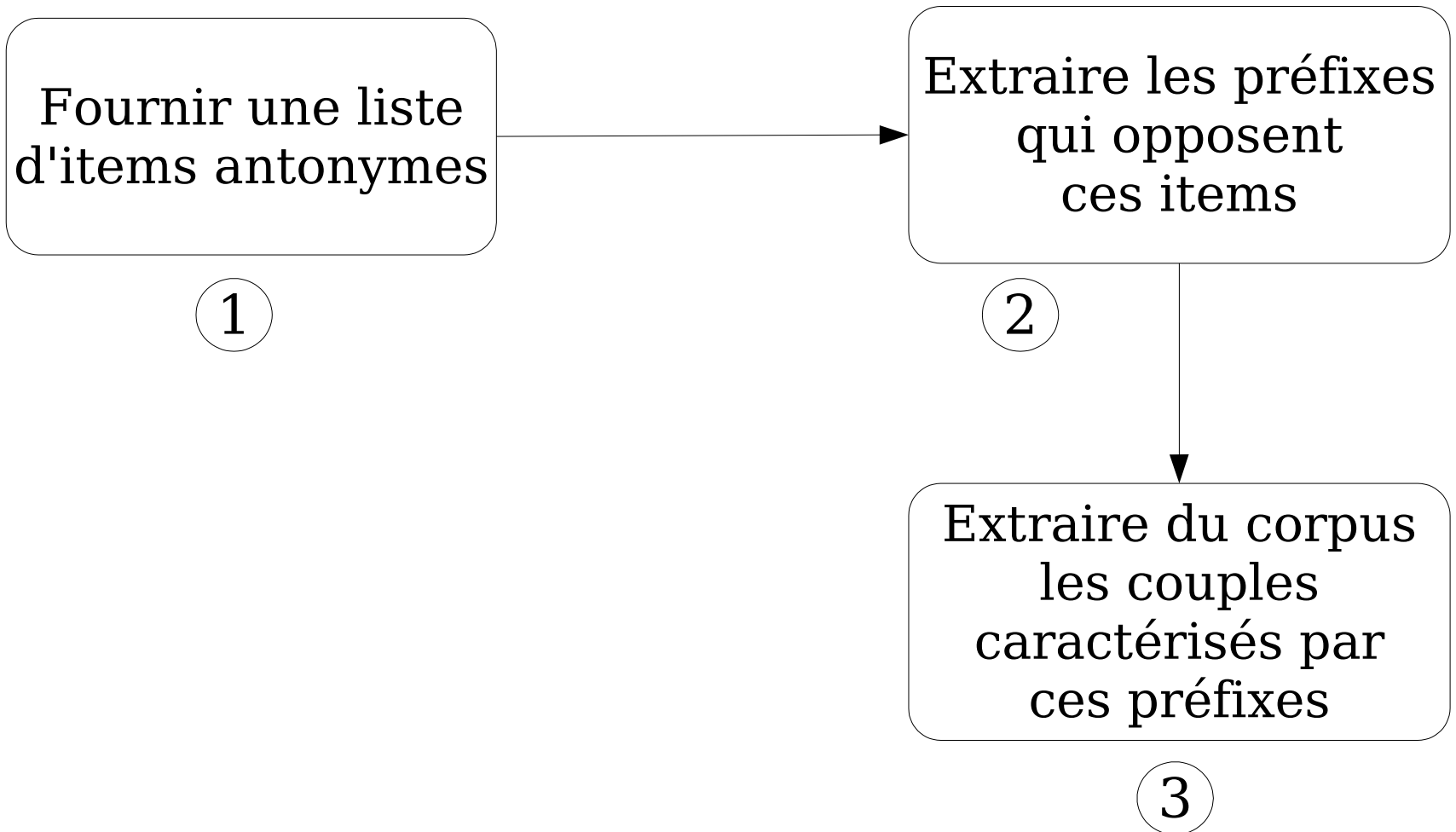
Fournir une liste
d'items antonymes

1

Extraire les préfixes
qui opposent
ces items

2

Processus d'acquisition





Processus d'acquisition

Extraire du corpus
les couples
caractérisés par
ces préfixes

3

Processus d'acquisition

Extraire du corpus
les couples
caractérisés par
ces préfixes

3

Valider ces couples
d'items et, ainsi,
les couples de
préfixes qui les
caractérisent

4

Processus d'acquisition

Extraire du corpus
les couples
caractérisés par
ces préfixes

3

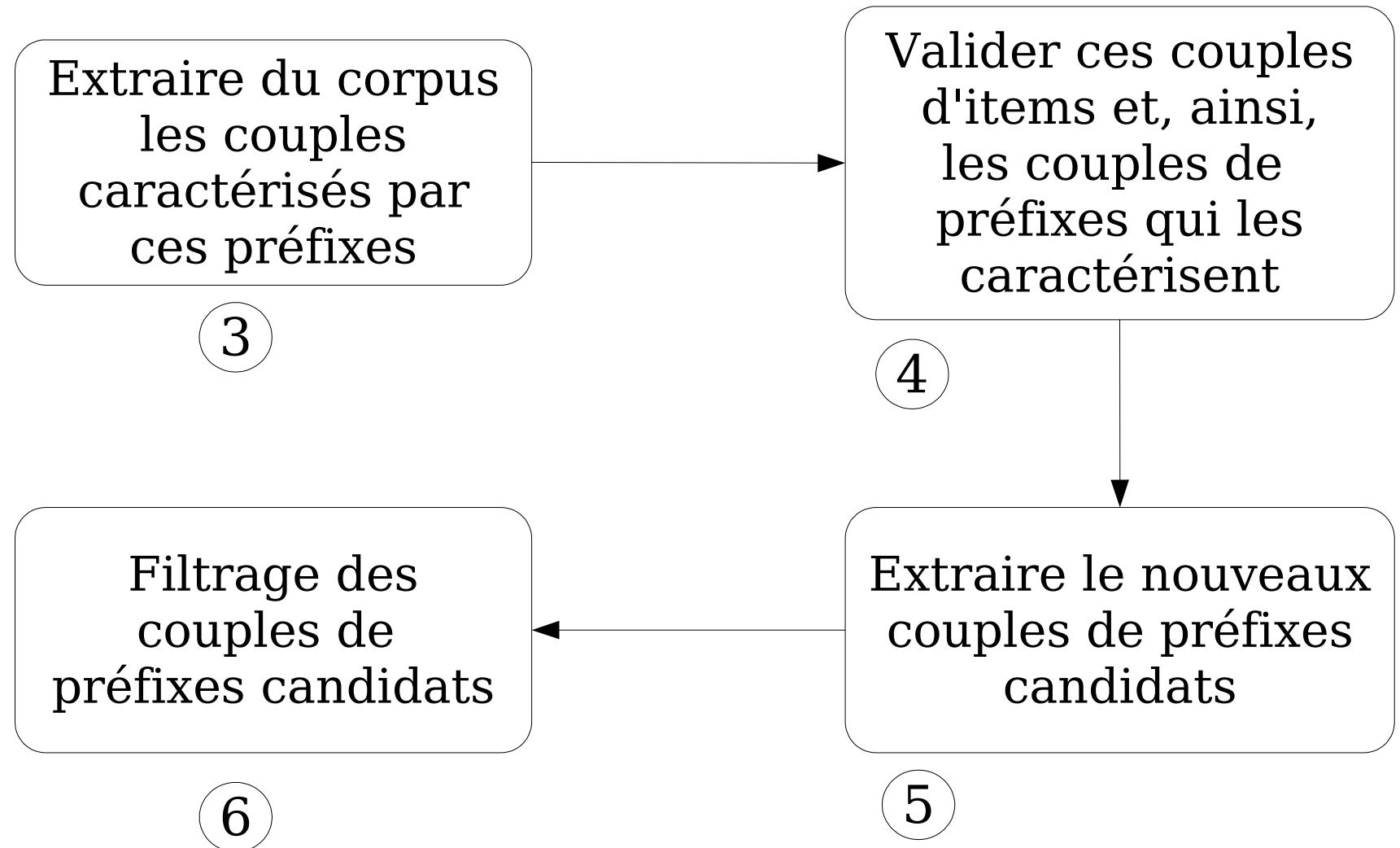
Valider ces couples
d'items et, ainsi,
les couples de
préfixes qui les
caractérisent

4

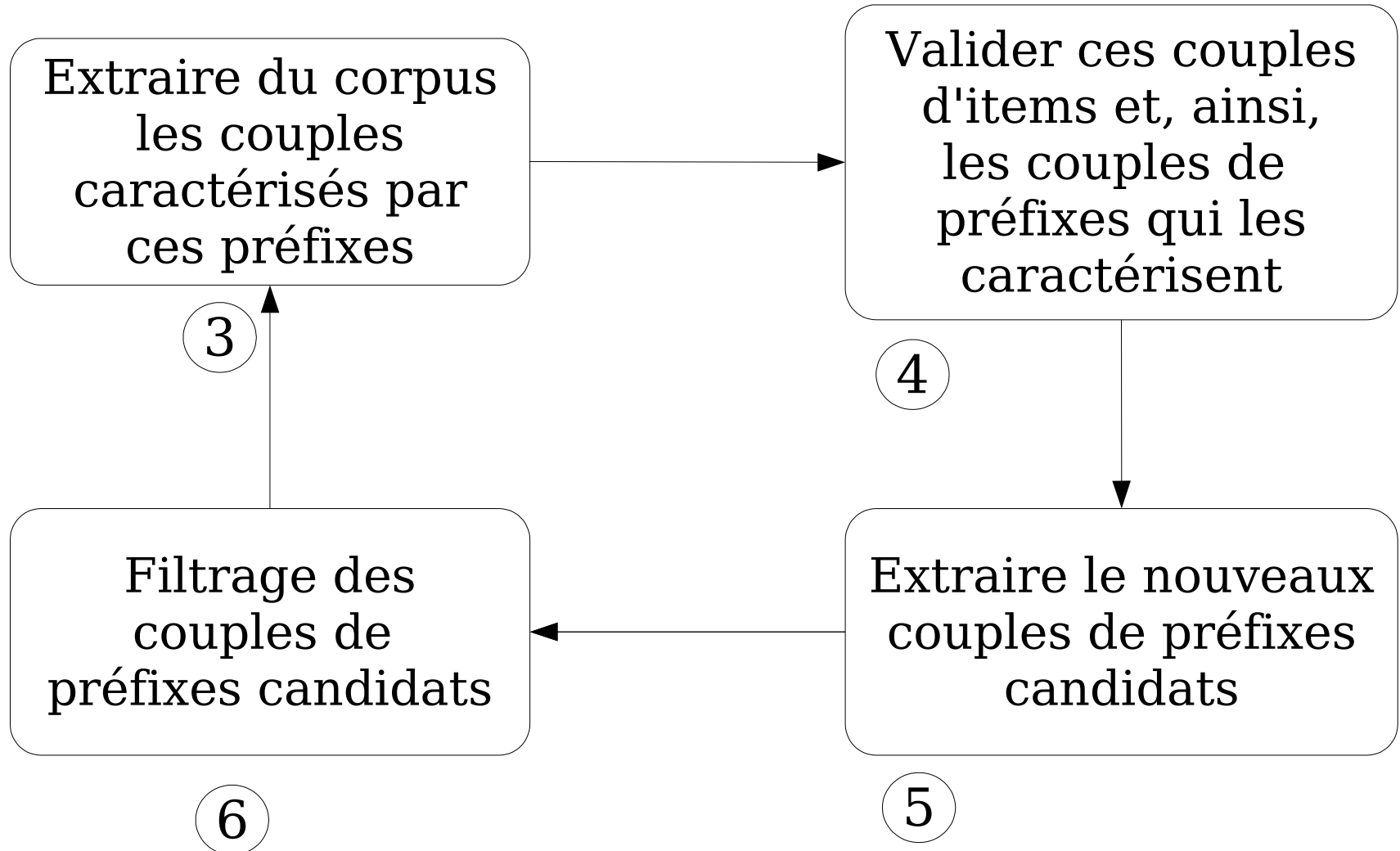
Extraire le nouveaux
couples de préfixes
candidats

5

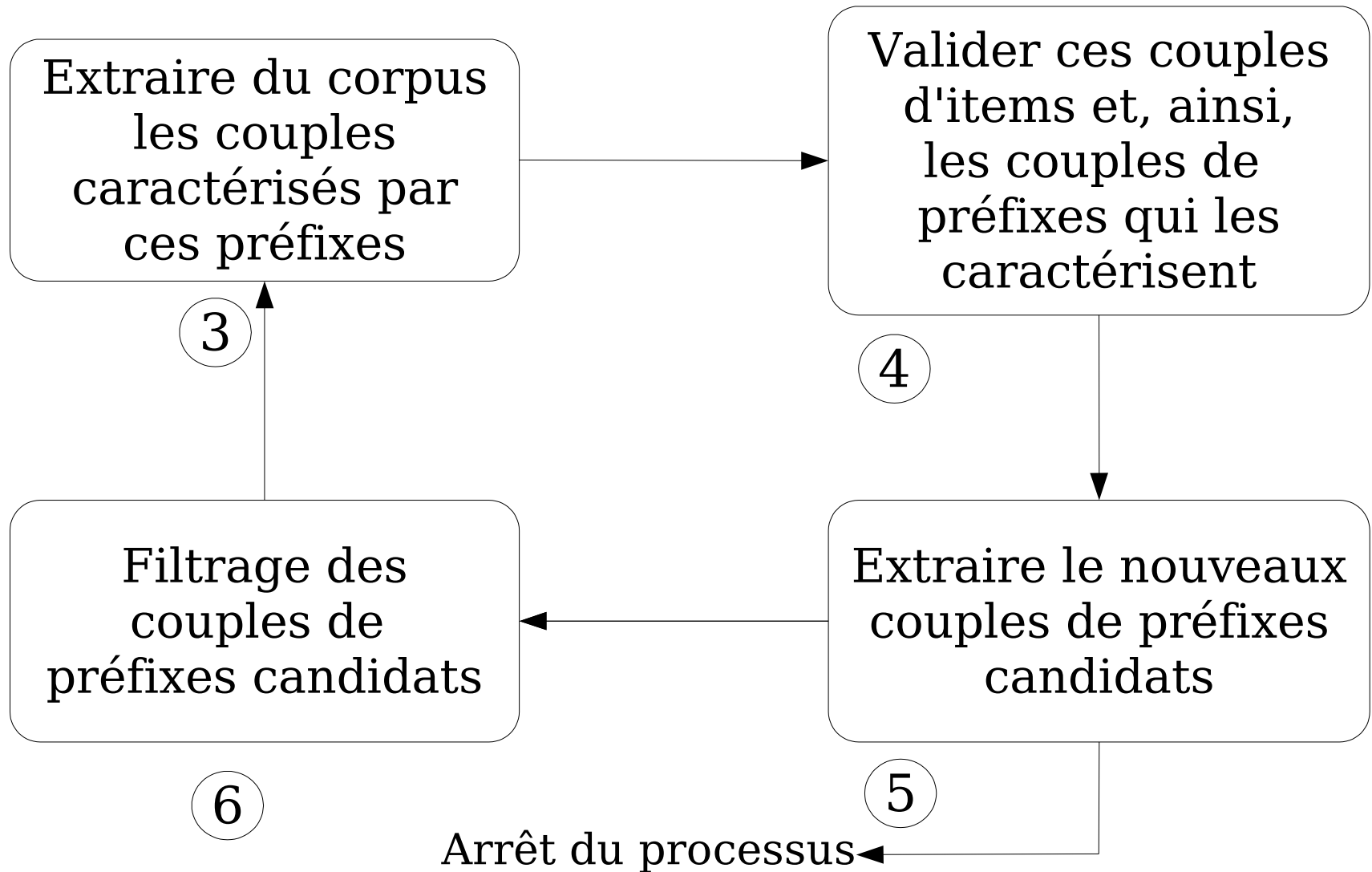
Processus d'acquisition



Processus d'acquisition



Processus d'acquisition





Déroulement

Corpus : *acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris*



Déroulement

1 - fournir une liste de couples d'items lexicaux antonymes

monosémique /// polysémique, acyclique /// cyclique



Déroulement

1 - fournir une liste de couples d'items lexicaux antonymes

monosémique /// polysémique, acyclique /// cyclique

2 - extraire les préfixes par lesquels ces couples s'opposent



Déroulement

1 - fournir une liste de couples d'items lexicaux antonymes

monosémique /// *polysémique*, *acyclique* /// *εcyclique*

2 - extraire les préfixes par lesquels ces couples s'opposent

mono- /// *poly-*



Déroulement

1 - fournir une liste de couples d'items lexicaux antonymes

monosémique /// *polysémique*, *a*cyclique /// *ε*cyclique

2 - extraire les préfixes par lesquels ces couples s'opposent

mono- /// *poly-*, *a-* /// *ε*



Déroulement

3 – extraire les préfixes par lesquels ces couples s'opposent

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris



Déroulement

3 – extraire les préfixes par lequel ces couples s'opposent

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

pour *mono-* /// *poly-* :



Déroulement

3 – extraire les préfixes par lequel ces couples s'opposent

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

pour mono- /// poly- : monosémique /// polysémique



Déroulement

3 – extraire les préfixes par lequel ces couples s'opposent

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

pour mono- /// poly- : monosémique /// polysémique



Déroulement

3 – extraire les préfixes par lequel ces couples s'opposent

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

*pour mono- /// poly- : monosémique /// polysémique
monophonie /// polyphonie*



Déroulement

3 – extraire les préfixes par lequel ces couples s'opposent

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

pour *mono-* /// *poly-* : *monosémique* /// *polysémique*
monophonie /// *polyphonie*

pour *a-* /// ϵ :



Déroulement

3 – extraire les préfixes par lequel ces couples s'opposent

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

pour *mono-* ||| *poly-* : *monosémique* ||| *polysémique*
monophonie ||| *polyphonie*

pour *a-* ||| ϵ : *acyclique* ||| *cyclique*



Déroulement

4 – validation par l'expert

pour *mono-* /// *poly-* : *monosémique* /// *polysémique*
monophonie /// *polyphonie*

pour *a-* /// ϵ : *acyclique* /// *cyclique*



Déroulement

4 – validation par l'expert

pour *mono-* /// *poly-* : *monosémique* /// *polysémique*
monophonie /// *polyphonie*

pour *a-* /// ϵ : *acyclique* ||| *cyclique*

monophonie : technique de reproduction des sons enregistrés ou transmis par radio, au moyen d'une seule voix.

polyphonie : chant à plusieurs voix.



Déroulement

4 – validation par l'expert

pour *mono-* ||| *poly-* : *monosémique* ||| *polysémique*
~~*monophonie* ||| *polyphonie*~~

pour *a-* ||| ϵ : *acyclique* ||| *cyclique*



Déroulement

4 – validation par l'expert

pour *mono-* ||| *poly-* : *monosémique* ||| *polysémique*
~~*monophonie* ||| *polyphonie*~~

pour *a-* ||| ϵ : *acyclique* ||| *cyclique*

mono- ||| *poly-* et *a-* ||| ϵ sont validés par l'expert



Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

5 – extraire de nouveaux préfixes candidats

monosémique ||| polysémique



Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

5 – extraire de nouveaux préfixes candidats

monosémique /// *polysémique*



Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

5 – extraire de nouveaux préfixes candidats

monosémique /// *polysémique*



Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

5 – extraire de nouveaux préfixes candidats

monosémique /// *polysémique*

mono- /// *bi-*, *poly-* /// *bi-*

Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

5 – extraire de nouveaux préfixes candidats

monosémique ||| polysémique

mono- ||| bi-, poly- ||| bi-

acyclique ||| cyclique :



Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

5 – extraire de nouveaux préfixes candidats

monosémique ||| polysémique

mono- ||| bi-, poly- ||| bi-

acyclique ||| cyclique :



Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

5 – extraire de nouveaux préfixes candidats

monosémique ||| polysémique

mono- ||| bi-, poly- ||| bi-

acyclique ||| cyclique :



Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, εcyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

5 – extraire de nouveaux préfixes candidats

monosémique ||| polysémique

mono- ||| bi-, poly- ||| bi-

acyclique ||| cyclique :

anti- ||| ε, a- ||| anti-



Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

6 – Filtrage des couples de préfixes candidats (n=2)

mono- ||| bi- : biphonie ||| monophonie, monosémique ||| bisémique



Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

6 – Filtrage des couples de préfixes candidats (n=2)

mono- /// bi- : biphonie /// monophonie, monosémique /// bisémique

poly- /// bi- : polyphonie /// biphonie



Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, ϵ cyclique, ϵ yclone, ϵ moine, monophonie, monosémique, polyphonie, polysémique, souris

6 – Filtrage des couples de préfixes candidats (n=2)

mono- /// bi- : biphonie /// monophonie, monosémique /// bisémique

poly- /// bi- : polyphonie /// biphonie

anti- /// ϵ : cyclone /// anticyclone, cyclique /// anticyclique, moine /// antimoine



Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

6 – Filtrage des couples de préfixes candidats (n=2)

mono- /// bi- : biphonie /// monophonie, monosémique /// bisémique

poly- /// bi- : polyphonie /// biphonie

anti- /// ε : cyclone /// anticyclone, cyclique /// anticyclique, moine /// antimoine

a- /// anti- : acyclique /// anticyclique



Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

6 – Filtrage des couples de préfixes candidats (n=2)

mono- /// bi- : biphonie /// monophonie, monosémique /// bisémique

~~*poly- /// bi- : polyphonie /// biphonie*~~

*anti- /// ε : cyclone /// anticyclone, cyclique /// anticyclique,
moine /// antimoine*

~~*a- /// anti- : acyclique /// anticyclique*~~



Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

7 – Validation des paires extraites par l'expert

mono- /// bi- : biphonie /// monophonie, monosémique /// bisémique

anti- /// ε : cyclone /// anticyclone, cyclique /// anticyclique, moine /// antimoine



Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

7 – Validation des paires extraites par l'expert

mono- /// bi- : biphonie /// monophonie, monosémique /// bisémique

*anti- /// ε : cyclone /// anticyclone, cyclique /// anticyclique,
~~moine /// antimoine~~*



Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

7 – Validation des paires extraites par l'expert

mono- /// bi- : biphonie /// monophonie, monosémique /// bisémique

anti- /// ε : cyclone /// anticyclone, cyclique /// anticyclique,

mono- /// bi- et anti- /// ε sont validés par l'expert



Déroulement

acyclique, anticyclique, anticyclone, antimoine, bisémique, biphonie, cyclique, cyclone, moine, monophonie, monosémique, polyphonie, polysémique, souris

Il n'y a plus de nouveaux préfixes candidats : fin du processus



Expérience menée

Corpus : liste d'items lexicaux (79220, issus de notre base)

Expert : Moi

Durée de l'expérience : Cinquantaine d'heures dont 99% pour la phase de validation (étape 4)

Résultats :

- Environ 2000 couples d'items antonymes extraits
- 59 couples de préfixes opposés extraits



Résultats (1/3)

a- ε	<i>chromatique</i> <i>achromatique</i>	<i>afin</i> <i>fin</i>
an- ε	<i>aérobie</i> <i>anaérobie</i>	<i>anatomiste</i> <i>atomiste</i>
anti- ε	<i>communiste-</i> <i>anticommuniste</i>	<i>moine</i> <i>antimoine</i>
dés- ε	<i>accord</i> <i>désaccord</i>	<i>avouer</i> <i>désavouer</i>
in- ε	<i>imaginable</i> <i>inimaginable</i>	<i>incas</i> <i>cas</i>
il- ε	<i>licite</i> <i>illicite</i>	<i>illustre</i> <i>lustre</i>
pré- post-	<i>préface-</i> <i>postface</i>	

Résultats (2/3)

	Paires extraites	Paires validées	Paires invalidées	taux validation
a- ε	288	71	217	24,6%
an- ε	61	15	41	24,6%
anti- ε	156	147	9	94,2%
dés- ε	195	179	16	91,7%
in- ε	616	539	77	87,5%
il- ε	31	18	13	58%
pré- post-	11	11	0	100%



Résultats(3/3)

Paires	Complémentaires	Scalaire	duals
a- ε	71 (100 %)		
an- ε	14 (92,8%)		1 (7,2%)
anti- ε			147 (100%)
dés- ε		3 (1,6 %)	176 (98,3%)
in- ε	539 (100 %)		
il- ε	18 (100%)		
pré- post-		11 (100%)	



Morphologie et type d'antonymie

La morphologie peut permettre globalement de connaître le type d'antonymie.

médicaux (hyper- ||| hypo-) : mesures au-dessus ou au-dessous de la normale (hyperthyroïdie ||| hypothyroïdie). scalaires

nombre (bi- ||| ε, tri-||| ε, mono- ||| poly-, mono- ||| stéréo-) : propriété possédée une fois (mono) ou plusieurs (bi-, tri-, poly, stéréo-). complémentaires.

absence de propriété : il- ||| ε (illicite ||| licite, illimité ||| limité), a- ||| ε, typique ||| atypique, sociabilité ||| associabilité, an- ||| ε, anencéphale ||| encéphale) complémentaires (exception : anion ||| ion duale).

contre : anti- ||| ε (anticléricale ||| cléricale, antiviral ||| viral). duals.



Conclusions

Méthode semi-supervisée permettant de construire le plus efficacement possible trois listes de couples d'antonymes.

Méthode basée sur les oppositions de nature morphologique
(*mono- /// poly-*, *a- /// ε*, *mono- /// stéréo*)

La même méthode peut être utilisée avec les suffixes ou d'autres relations sémantiques

suffixes : *-phile /// -phobe*; *hydrophobe /// hydrophile*

synonymie : *di- ≡ bi-*; *dipolaire ≡ bipolaire*

dis- ≡ in-; *disharmonie ≡ inharmonie*