

# Une métrologie robuste des réseaux lexicaux fondée sur la dynamique des balades aléatoires courtes à travers les relations lexicales

Bruno Gaume

CLLE, CNRS & Univ. Toulouse

Réseaux Sémantiques : Construction, Exploration et Utilisations

Marseille, 12 Juin 2017

- 1 Structure des réseaux lexicaux
- 2 Mesurer la similarité structurelle entre réseaux lexicaux

## **Applications**

- 3 WISIGOTH : Enrichissement de lexiques approvisionnés par les foules
- 4 SLAM : Substitution lexicale en contexte
- 5 CELLO : Détection de communautés pour la Recherche d'Information
- 6 SKILLEX : Un score d'habileté lexicale

## 1 Structure des réseaux lexicaux

2 Mesurer la similarité structurelle entre réseaux lexicaux

### Applications

3 WISIGOTH : Enrichissement de lexiques approvisionnés par les foules

4 SLAM : Substitution lexicale en contexte

5 CELLO : Détection de communautés pour la Recherche d'Information

6 SKILLEX : Un score d'habileté lexicale

- ▶ Gaume B. (2004). Balades aléatoires dans les petits mondes lexicaux. *Information Interaction Intelligence*, 4(2).
- ▶ Gaume B. (2008) Mapping the form of meaning in small worlds. *Journal of Intelligent Systems*, 23(7) :848862.
- ▶ Gaume B., Mathieu F., and Navarro E. (2010). Building real-world complex networks by wandering on random graphs. *Information Interaction Intelligence*, 10(1).

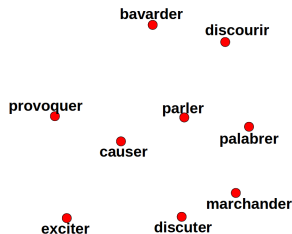
# Structure des réseaux lexicaux

$\mathbf{G} = (\mathbf{V}, \mathbf{E})$  où  $E \subseteq P_2^V$  : un graphe de  $n = |V|$  sommets et  $m = |E|$  arêtes

# Structure des réseaux lexicaux

$\mathbf{G} = (\mathbf{V}, \mathbf{E})$  où  $E \subseteq P_2^V$  : un graphe de  $n = |V|$  sommets et  $m = |E|$  arêtes

$\mathbf{V} = \{\text{bavarder, causer, discourir, discuter, exciter, palabrer, parler, pérorer, provoquer}\}$

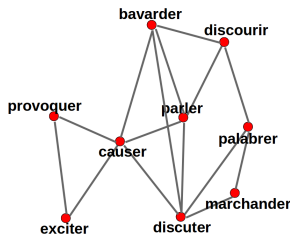


# Structure des réseaux lexicaux

$G = (V, E)$  où  $E \subseteq P_2^V$  : un graphe de  $n = |V|$  sommets et  $m = |E|$  arêtes

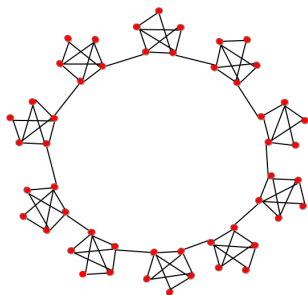
$V = \{\text{bavarder, causer, discourir, discuter, exciter, palabrer, parler, p rer, provoquer}\}$

$E = \{\{\text{bavarder, parler}\}, \{\text{bavarder, discourir}\}, \{\text{bavarder, causer}\}, \{\text{causer, discuter}\}, \{\text{causer, parler}\}, \{\text{causer, provoquer}\}, \{\text{causer, exciter}\}, \{\text{discourir, parler}\}, \{\text{discourir, palabrer}\}, \{\text{discourir, p rer}\}, \{\text{discuter, parler}\}, \{\text{discuter, palabrer}\}, \{\text{discuter, p rer}\}, \{\text{exciter, provoquer}\}, \{\text{palabrer, parler}\}, \{\text{palabrer, p rer}\}, \{\text{parler, p rer}\}\}$



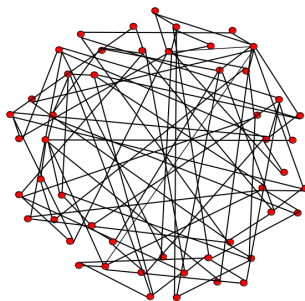
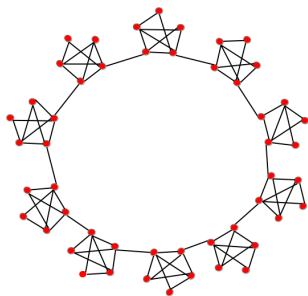
## Graphes de laboratoire

## Graphes de laboratoire





## Graphes de laboratoire



## Graphes de terrain

## Graphes de terrain

- ▶ Les graphes d'accointance d'un groupe d'humains
- ▶ Le graphe des appels téléphoniques aux USA
- ▶ Le graphe du World Wide Web
- ▶ Le graphe des collaborations cinématographiques
- ▶ Le graphe de *Caenorhabditis elegans*
- ▶ Les graphes Lexicaux

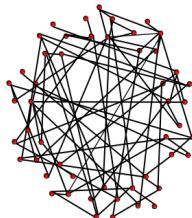
## Graphes de terrain

- ▶ Les graphes d'accointance d'un groupe d'humains
- ▶ Le graphe des appels téléphoniques aux USA
- ▶ Le graphe du World Wide Web
- ▶ Le graphe des collaborations cinématographiques
- ▶ Le graphe de *Caenorhabditis elegans*
- ▶ Les graphes Lexicaux

## Réseaux Petits Mondes Hiérarchiques

- ▶ Faible densité
- ▶ Chemins courts
- ▶ Distribution des degrés à queue lourde (loi de puissance)
- ▶ Fort coefficient de clustering : zones denses en arêtes

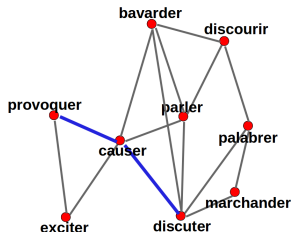
## Graphes de terrain



## Réseaux Petits Mondes Hiérarchiques

- ▶ **Faible densité**
- ▶ Chemins courts
- ▶ Distribution des degrés à queue lourde (loi de puissance)
- ▶ Fort coefficient de clustering : zones denses en arêtes

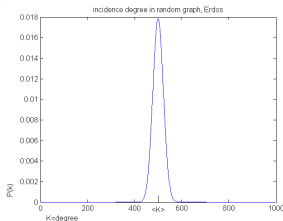
## Graphes de terrain



## Réseaux Petits Mondes Hiérarchiques

- ▶ Faible densité
- ▶ **Chemins courts**
- ▶ Distribution des degrés à queue lourde (loi de puissance)
- ▶ Fort coefficient de clustering : zones denses en arêtes

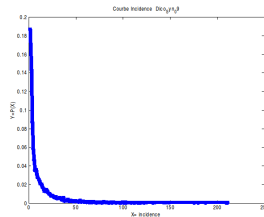
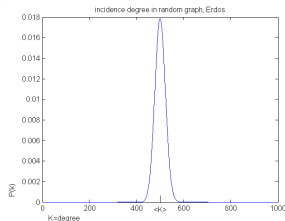
## Graphes de terrain



## Réseaux Petits Mondes Hiérarchiques

- ▶ Faible densité
- ▶ Chemins courts
- ▶ **Distribution des degrés à queue lourde (loi de puissance)**
- ▶ Fort coefficient de clustering : zones denses en arêtes

## Graphes de terrain

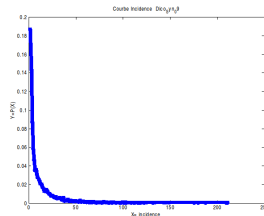
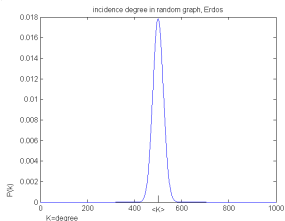


## Réseaux Petits Mondes Hiérarchiques

- ▶ Faible densité
- ▶ Chemins courts
- ▶ **Distribution des degrés à queue lourde (loi de puissance)**
- ▶ Fort coefficient de clustering : zones denses en arêtes



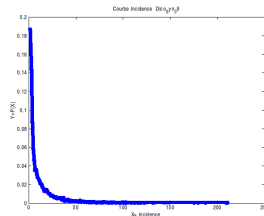
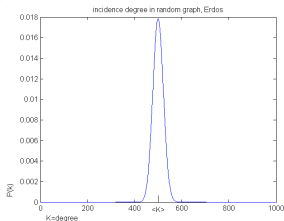
## Graphes de terrain



## Réseaux Petits Mondes Hiérarchiques

- ▶ Faible densité
- ▶ Chemins courts
- ▶ **Distribution des degrés à queue lourde (loi de puissance)**
- ▶ Fort coefficient de clustering : zones denses en arêtes

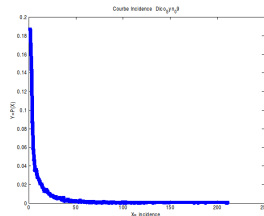
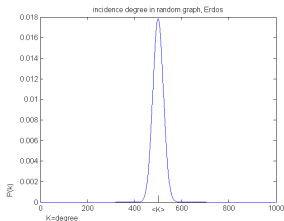
## Graphes de terrain



## Réseaux Petits Mondes Hiérarchiques

- ▶ Faible densité
- ▶ Chemins courts
- ▶ **Distribution des degrés à queue lourde (loi de puissance)**
- ▶ Fort coefficient de clustering : zones denses en arêtes

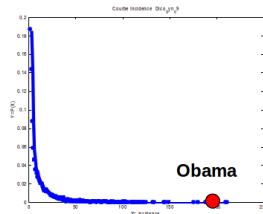
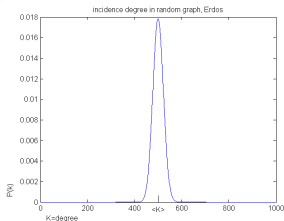
## Graphes de terrain



## Réseaux Petits Mondes Hiérarchiques

- ▶ Faible densité
- ▶ Chemins courts
- ▶ **Distribution des degrés à queue lourde (loi de puissance)**
- ▶ Fort coefficient de clustering : zones denses en arêtes

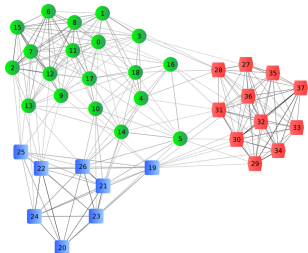
## Graphes de terrain



## Réseaux Petits Mondes Hiérarchiques

- ▶ Faible densité
- ▶ Chemins courts
- ▶ **Distribution des degrés à queue lourde (loi de puissance)**
- ▶ Fort coefficient de clustering : zones denses en arêtes

## Graphes de terrain



## Réseaux Petits Mondes Hiérarchiques

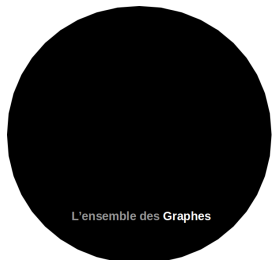
- ▶ Faible densité
- ▶ Chemins courts
- ▶ Distribution des degrés à queue lourde (loi de puissance)
- ▶ **Fort coefficient de clustering : zones denses en arêtes**

## Graphes de terrain

### Réseaux Petits Mondes Hiérarchiques

- ▶ Faible densité
- ▶ Chemins courts
- ▶ Distribution des degrés à queue lourde (loi de puissance)
- ▶ Fort coefficient de clustering : zones denses en arêtes

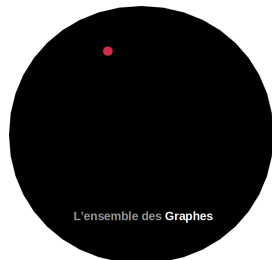
## Graphes de terrain



### Réseaux Petits Mondes Hiérarchiques

- ▶ Faible densité
- ▶ Chemins courts
- ▶ Distribution des degrés à queue lourde (loi de puissance)
- ▶ Fort coefficient de clustering : zones denses en arêtes

## Graphes de terrain

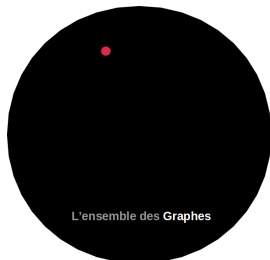


### Réseaux Petits Mondes Hiérarchiques

- ▶ Faible densité
- ▶ Chemins courts
- ▶ Distribution des degrés à queue lourde (loi de puissance)
- ▶ Fort coefficient de clustering : zones denses en arêtes



## Graphes de terrain

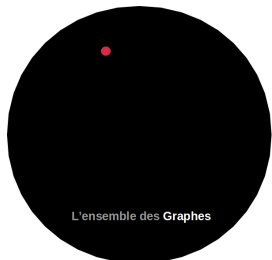


- Les graphes d'accointance d'un groupe d'humains
- Le graphe des appels téléphoniques aux USA
- Le graphe du World Wide Web
- Le graphe des collaborations cinématographiques
- Le graphe de *Caenorhabditis elegans*
- **Les graphes Lexicaux**

## Réseaux Petits Mondes Hiérarchiques

- ▶ Faible densité
- ▶ Chemins courts
- ▶ Distribution des degrés à queue lourde (loi de puissance)
- ▶ Fort coefficient de clustering : zones denses en arêtes

## Graphes de terrain



- Les graphes d'accointance d'un groupe d'humains
- Le graphe des appels téléphoniques aux USA
- Le graphe du World Wide Web
- Le graphe des collaborations cinématographiques
- Le graphe de *Caenorhabditis elegans*
- **Les graphes Lexicaux**

## Réseaux Petits Mondes Hiérarchiques

- ▶ Faible densité
- ▶ Chemins courts
- ▶ **Distribution des degrés à queue lourde (loi de puissance)**
- ▶ **Fort coefficient de clustering : zones denses en arêtes**

1 Structure des réseaux lexicaux

2 **Mesurer la similarité structurelle entre réseaux lexicaux**

## Applications

3 WISIGOTH : Enrichissement de lexiques approvisionnés par les foules

4 SLAM : Substitution lexicale en contexte

5 CELLO : Détection de communautés pour la Recherche d'Information

6 SKILLEX : Un score d'habileté lexicale

► Gaillard B., Gaume B., and Navarro E. (2011). Invariants and variability of synonymy networks : Self mediated agreement **in** Proc. of the TextGraphs-6 2011 : Graph-based Methods for Natural Language Processing, Portland, Oregon, USA.

► Gaume B., Navarro E., Desalle Y., Gaillard B. (2014). Mesurer la similarité structurelle entre réseaux lexicaux **in** Proc. of TALN-20 2014 : Atelier RLTLN, Réseaux Lexicaux et Traitement des Langues Naturelles, Marseille, France.

# Problème

**Hypothèse** : La synonymie reflète la structure sémantique du lexique commune aux membres d'une même communauté linguistique

**Hypothèse** : La synonymie reflète la structure sémantique du lexique commune aux membres d'une même communauté linguistique

**Incompatible avec :**

**Constat** : Il existe de nombreux désaccords entre les membres d'une même communauté linguistique :

- ▶ **Beaucoup** de paires de mots sont synonymes pour l'individu  $I_1$  et pas pour l'individu  $I_2$  et vice et versa



**Hypothèse** : La synonymie reflète la structure sémantique du lexique commune aux membres d'une même communauté linguistique

## Résoudre cette contradiction apparente

**Constat** : Il existe de nombreux désaccords entre les membres d'une même communauté linguistique :

- ▶ **Beaucoup** de paires de mots sont synonymes pour l'individu  $I_1$  et pas pour l'individu  $I_2$  et vice et versa







**bavarder** : parler, discourir, causer

**causer** : parler, discuter, bavarder, provoquer, exciter

**discourir** : parler, palabrer, pérorer, bavarder

**discuter** : parler, palabrer, causer, pérorer

**exciter** : causer, provoquer

**palabrer** : parler, discourir, pérorer, discuter

**parler** : palabrer, discourir, causer, pérorer, discuter, bavarder

**pérorer** : parler, palabrer, discourir, discuter

**provoquer** : causer, exciter

**bavarder** : parler, discourir, causer

**causer** : parler, discuter, bavarder, provoquer, exciter

**discourir** : parler, palabrer, pérorer, bavarder

**discuter** : parler, palabrer, causer, pérorer

**exciter** : causer, provoquer

**palabrer** : parler, discourir, pérorer, discuter

**parler** : palabrer, discourir, causer, pérorer, discuter, bavarder

**pérorer** : parler, palabrer, discourir, discuter

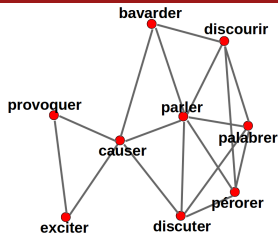
**provoquer** : causer, exciter

$G = (V, E)$  où  $E \subseteq P_2^V$

$V = \{\text{bavarder, causer, discourir, discuter, exciter, palabrer, parler, pérorer, provoquer}\}$

$E = \{\{\text{bavarder, parler}\}, \{\text{bavarder, discourir}\}, \{\text{bavarder, causer}\}, \{\text{causer, discuter}\}, \{\text{causer, parler}\}, \{\text{causer, provoquer}\}, \{\text{causer, exciter}\}, \{\text{discourir, parler}\}, \{\text{discourir, palabrer}\}, \{\text{discourir, pérorer}\}, \{\text{discuter, parler}\}, \{\text{discuter, palabrer}\}, \{\text{discuter, pérorer}\}, \{\text{exciter, provoquer}\}, \{\text{palabrer, parler}\}, \{\text{palabrer, pérorer}\}, \{\text{parler, pérorer}\}\}$

Un graphe de  $n = |V|$  sommets et  $m = |E|$  arêtes

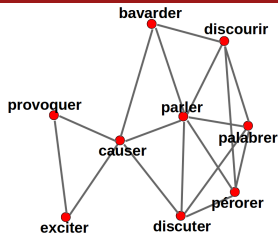


$G = (V, E)$  où  $E \subseteq P_2^V$

$V = \{\text{bavarder, causer, discourir, discuter, exciter, palabrer, parler, pérorer, provoquer}\}$

$E = \{\{\text{bavarder, parler}\}, \{\text{bavarder, discourir}\}, \{\text{bavarder, causer}\}, \{\text{causer, discuter}\}, \{\text{causer, parler}\}, \{\text{causer, provoquer}\}, \{\text{causer, exciter}\}, \{\text{discourir, parler}\}, \{\text{discourir, palabrer}\}, \{\text{discourir, pérorer}\}, \{\text{discuter, parler}\}, \{\text{discuter, palabrer}\}, \{\text{discuter, pérorer}\}, \{\text{exciter, provoquer}\}, \{\text{palabrer, parler}\}, \{\text{palabrer, pérorer}\}, \{\text{parler, pérorer}\}\}$

Un graphe de  $n = |V|$  sommets et  $m = |E|$  arêtes



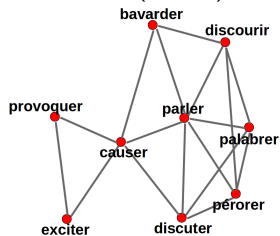
## Réseaux Petits Mondes Hiérarchiques

- ▶ Faible densité
- ▶ Chemins courts
- ▶ Distribution des degrés à queue lourde (loi de puissance)
- ▶ Fort coefficient de clustering : zones denses en arêtes

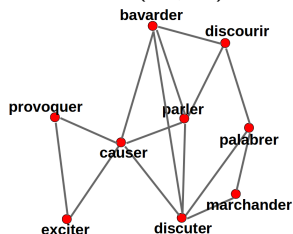
# Comparer $E_1$ et $E_2$ comme des «sacs de liens»

# Comparer $E_1$ et $E_2$ comme des «sacs de liens»

$$G_1 = (V_1, E_1)$$

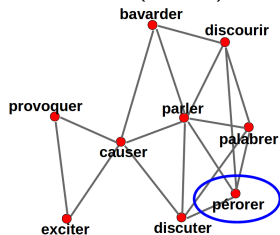


$$G_2 = (V_2, E_2)$$

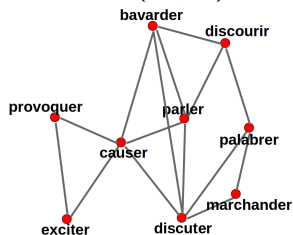


# Comparer $E_1$ et $E_2$ comme des «sacs de liens»

$$G_1 = (V_1, E_1)$$

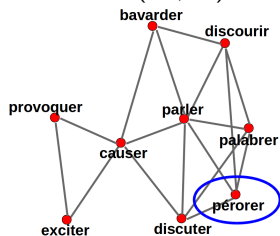


$$G_2 = (V_2, E_2)$$

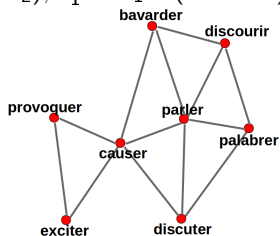


# Comparer $E_1$ et $E_2$ comme des «sacs de liens»

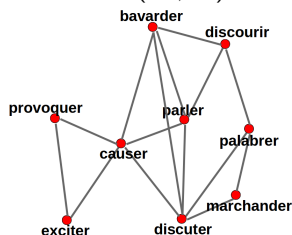
$$G_1 = (V_1, E_1)$$



$$G'_1 = (V' = (V_1 \cap V_2), E'_1 = E_1 \cap (V' \times V'))$$



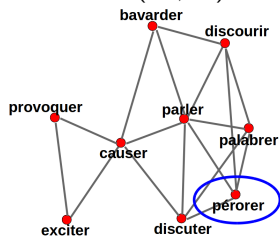
$$G_2 = (V_2, E_2)$$



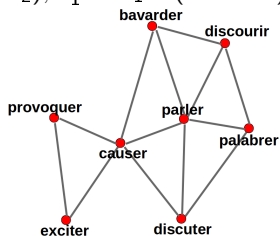


# Comparer $E_1$ et $E_2$ comme des «sacs de liens»

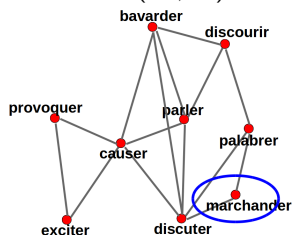
$$G_1 = (V_1, E_1)$$



$$G'_1 = (V' = (V_1 \cap V_2), E'_1 = E_1 \cap (V' \times V'))$$

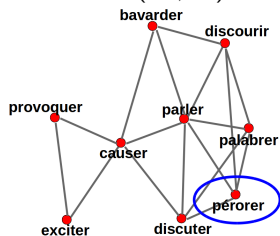


$$G_2 = (V_2, E_2)$$

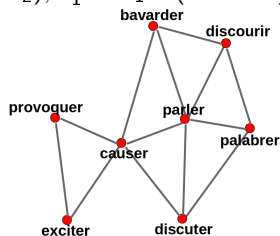


# Comparer $E_1$ et $E_2$ comme des «sacs de liens»

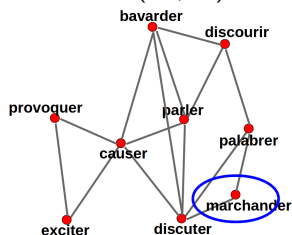
$$G_1 = (V_1, E_1)$$



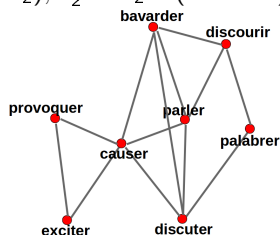
$$G'_1 = (V' = (V_1 \cap V_2), E'_1 = E_1 \cap (V' \times V'))$$



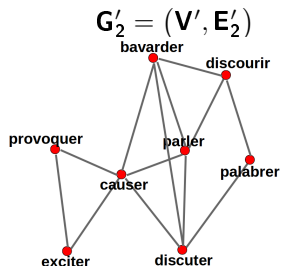
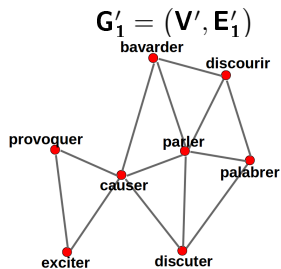
$$G_2 = (V_2, E_2)$$



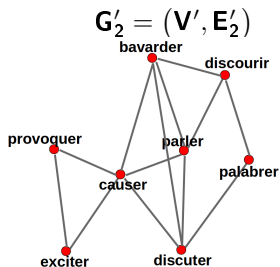
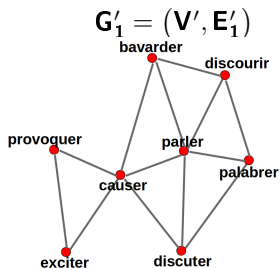
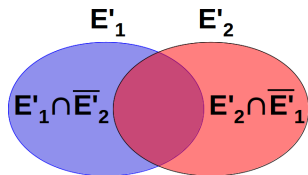
$$G'_2 = (V' = (V_1 \cap V_2), E'_2 = E_2 \cap (V' \times V'))$$



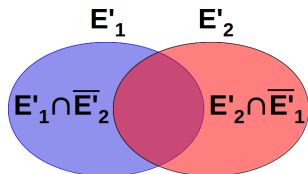
# Comparer $E_1$ et $E_2$ comme des «sacs de liens»



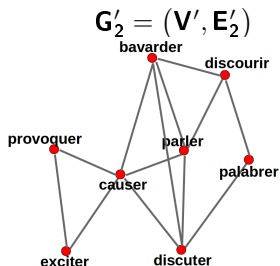
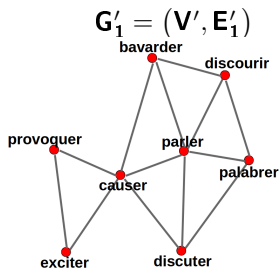
# Comparer $E_1$ et $E_2$ comme des «sacs de liens»



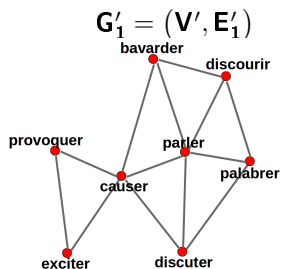
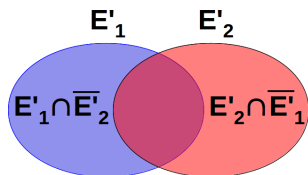
# Comparer $E_1$ et $E_2$ comme des «sacs de liens»



$$GED(G'_1, G'_2) = \frac{|E'_1 \cap \overline{E'_2}| + |E'_2 \cap \overline{E'_1}|}{|E'_1| + |E'_2|} \in [0, 1]$$

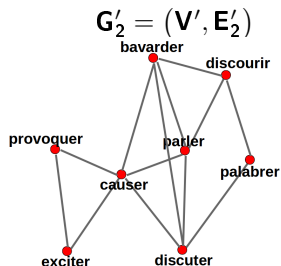


# Comparer $E_1$ et $E_2$ comme des «sacs de liens»

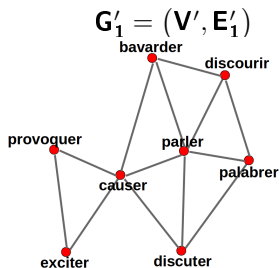
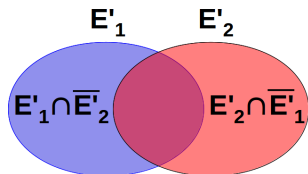


$$GED(G'_1, G'_2) = \frac{|E'_1 \cap \overline{E'_2}| + |E'_2 \cap \overline{E'_1}|}{|E'_1| + |E'_2|} \in [0, 1]$$

$$GED(Lar', Rob') = 0.47$$



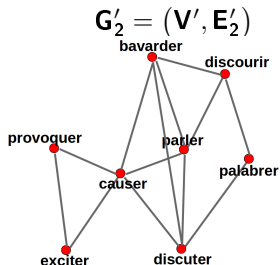
# Comparer $E_1$ et $E_2$ comme des «sacs de liens»



$$GED(G'_1, G'_2) = \frac{|E'_1 \cap \overline{E'_2}| + |E'_2 \cap \overline{E'_1}|}{|E'_1| + |E'_2|} \in [0, 1]$$

$$GED(Lar', Rob') = 0.47$$

Nombreux désaccords entre Larousse et Robert



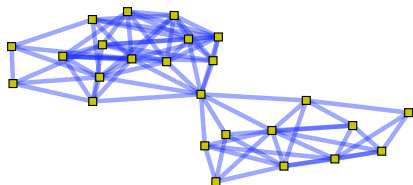
# Hypothèse



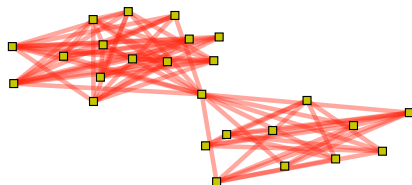
Nombreux désaccords **au niveau des arêtes**

Nombreux désaccords **au niveau des arêtes**

$G_1 = (V, E_1)$



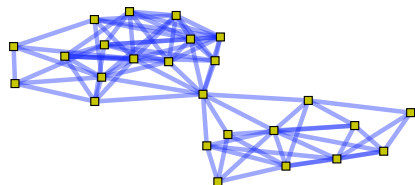
$G_2 = (V, E_2)$



**Aucune arête commune** entre ces deux graphes.

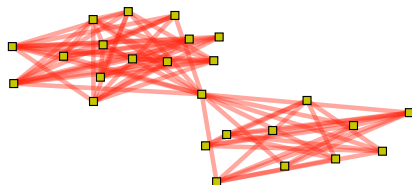
Nombreux désaccords **au niveau des arêtes**

$G_1 = (V, E_1)$



$\text{GED}(G_1, G_2) = 1$

$G_2 = (V, E_2)$



**Aucune arête commune** entre ces deux graphes.

Nombreux désaccords **au niveau des arêtes**

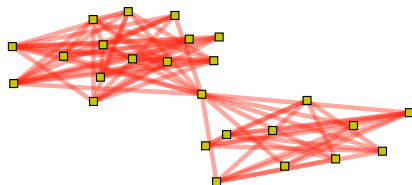
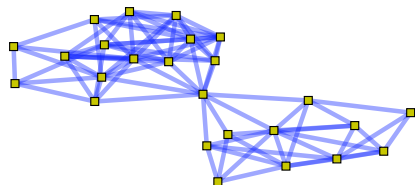
Mais

Bon accord **au niveau des clusters...**

$G_1 = (V, E_1)$

$\text{GED}(G_1, G_2) = 1$

$G_2 = (V, E_2)$



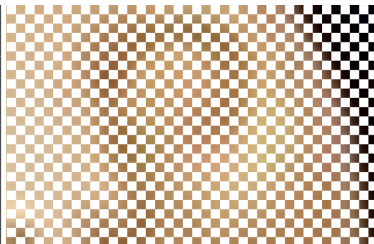
**Aucune arête commune** entre ces deux graphes.

# Métaphore

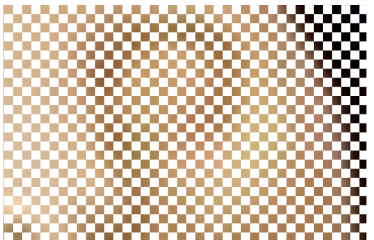
# Métaphore



La photo d'un chat



Chaque pixel **impair** est peint en blanc



Chaque pixel **pair** est peint en blanc

# Métaphore

*Pair =*

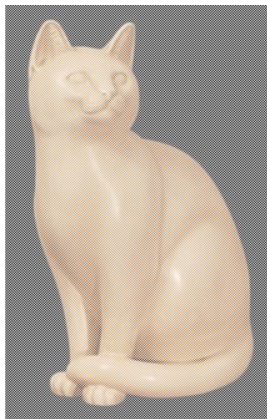


# Métaphore

*Pair* =



*Impair* =





# Métaphore

*Pair* =



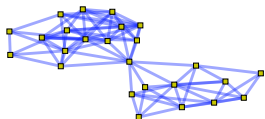
*Impair* =



**Aucun pixel commun**  
**Images distantes (au niveau des pixels)**  
**Cependant ces deux images se ressemblent**

# Métaphore

*Pair* =



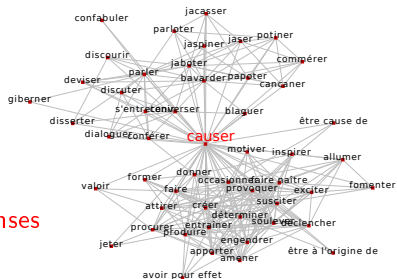
*Impair* =



# Marches Aléatoires

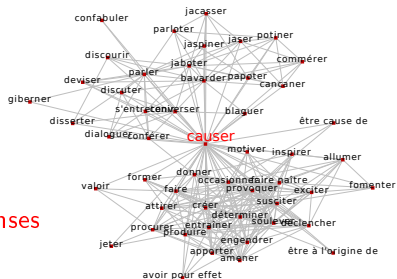
## Réseaux Petits Mondes Hiérarchiques

- ▶ Faible densité
- ▶ Chemins courts
- ▶ Distribution des degrés à queue lourde
- ▶ Fort coefficient de clustering : **zones denses**



## Réseaux Petits Mondes Hiérarchiques

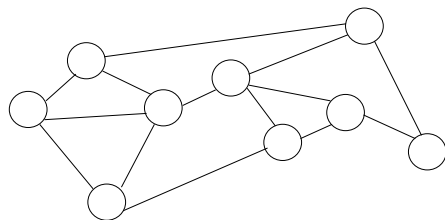
- ▶ Faible densité
- ▶ Chemins courts
- ▶ Distribution des degrés à queue lourde
- ▶ Fort coefficient de clustering : **zones denses**



## Marches Aléatoires

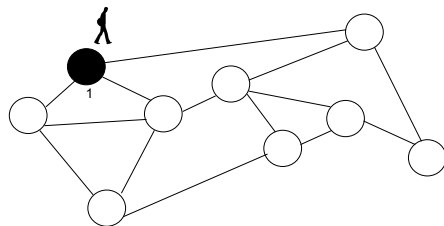
- ▶ **Idée**: si  $(u, v)$  sont **dans un même "cluster"**, alors **beaucoup de chemins** mènent de  $u$  à  $v$  (même si ils ne sont pas adjacents).
- ▶ Un **marcheur aléatoire** a tendance à être **capturé par les "clusters"**

# Marches Aléatoires



- 1 État initial ( $t=0$ )
- 2  $t = 1$
- 3  $t = 2$
- 4  $t = 3$

# Marches Aléatoires



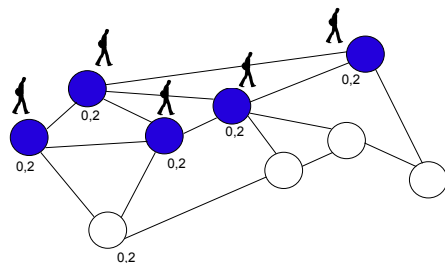
1 État initial ( $t=0$ )

2  $t = 1$

3  $t = 2$

4  $t = 3$

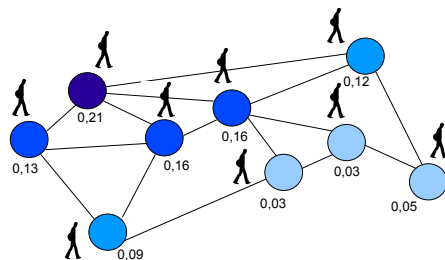
# Marches Aléatoires



- 1 État initial ( $t=0$ )
- 2  $t = 1$
- 3  $t = 2$
- 4  $t = 3$

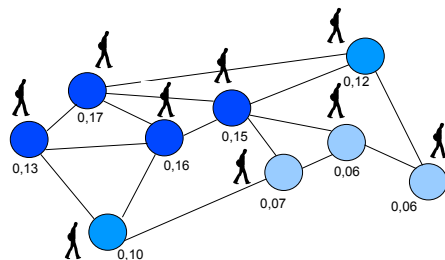


# Marches Aléatoires



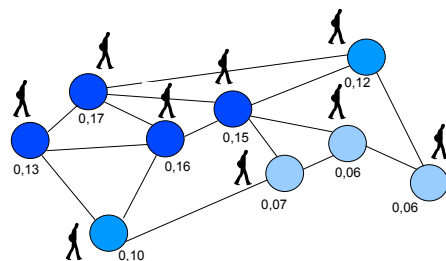
- 1 État initial ( $t=0$ )
- 2  $t = 1$
- 3  $t = 2$
- 4  $t = 3$

# Marches Aléatoires



- 1 État initial ( $t=0$ )
- 2  $t = 1$
- 3  $t = 2$
- 4  $t = 3$

# Marches Aléatoires



- 1 État initial ( $t=0$ )
- 2  $t = 1$
- 3  $t = 2$
- 4  $t = 3$

La Proxémie :  $P_G^t(u, v)$

Dans  $G$ , probabilité d'être sur  $v$  à l'instant  $t$  en partant de  $u$

# Marches Aléatoires

Liste ordonnées des verbes selon leur score de proxémie à partir de *déshabiller*  
( $t = 4$ )

1) dépouiller*	11) dévoiler	21) arracher	31) tailler	41) dire
2) défaire*	12) dégager	22) délacer*	32) développer	42) ouvrir
3) démunir*	13) étaler*	23) voler	33) déchirer	43) déployer
4) déshabiller*	14) ôter*	24) dénouer	34) déchausser	44) dépiauter
5) dévêtir*	15) enlever	25) desserrer	35) vider	45) déposséder
6) révéler	16) défrusquer	26) médire*	36) peler	46) frustrer
7) découvrir*	17) afficher	27) débarrasser	37) dépourvoir	47) indiquer
8) montrer*	18) désaffubler	28) exposer	38) défubler	48) démontrer
9) dénuder*	19) écorcher*	29) couper	39) priver	49) manifester
10) dégarnir*	20) démasquer*	30) exhiber*	40) trahir	50) détacher

**Table :** Les cinquante verbes qui ont la plus forte probabilité d'être atteints en 4 pas depuis *déshabiller* (sur 10000 verbes).

La Proxémie :  $P_G^t(u, v)$

Dans  $G$ , probabilité d'être sur  $v$  à l'instant  $t$  en partant de  $u$

# Marches Aléatoires

Liste ordonnées des verbes selon leur score de proxémie à partir de *déshabiller*  
( $t = 4$ )

1) dépouiller*	11) dévoiler	21) arracher	31) tailler	41) dire
2) défaire*	12) dégager	22) délacer*	32) développer	42) ouvrir
3) démunir*	13) étaler*	23) voler	33) déchirer	43) déployer
4) <b>déshabiller*</b>	14) ôter*	24) dénouer	34) déchausser	44) dépiauter
5) dévêtir*	15) enlever	25) desserrer	35) vider	45) déposséder
6) révéler	16) défrusquer	26) médire*	36) peler	46) frustrer
7) découvrir*	17) afficher	27) débarrasser	37) dépourvoir	47) indiquer
8) montrer*	18) désaffubler	28) exposer	38) défubler	48) démontrer
9) dénuder*	19) écorcher*	29) couper	39) priver	49) manifester
10) dégarnir*	20) démasquer*	30) exhiber*	40) trahir	50) détacher

**Table :** Les cinquante verbes qui ont la plus forte probabilité d'être atteints en 4 pas depuis *déshabiller* (sur 10000 verbes).

La Proxémie :  $P_G^t(u, v)$

Dans  $G$ , probabilité d'être sur  $v$  à l'instant  $t$  en partant de  $u$

# Marches Aléatoires

Liste ordonnées des verbes selon leur score de proxémie à partir de *déshabiller*  
( $t = 4$ )

1) dépouiller*	11) dévoiler	21) arracher	31) tailler	41) dire
2) défaire*	12) dégager	22) délacer*	32) développer	42) ouvrir
3) démunir*	13) étaler*	23) voler	33) déchirer	43) déployer
4) déshabiller*	14) ôter*	24) dénouer	34) déchausser	44) dépiauter
5) dévêtir*	15) enlever	25) desserrer	35) vider	45) déposséder
6) révéler	16) défrusquer	26) médire*	36) peler	46) frustrer
7) découvrir*	17) afficher	27) débarrasser	37) dépourvoir	47) indiquer
8) montrer*	18) désaffubler	28) exposer	38) défubler	48) démontrer
9) dénuder*	19) écorcher*	29) couper	39) priver	49) manifester
10) dégarnir*	20) démasquer*	30) exhiber*	40) trahir	50) détacher

**Table :** Les cinquante verbes qui ont la plus forte probabilité d'être atteints en 4 pas depuis *déshabiller* (sur 10000 verbes).

La Proxémie :  $P_G^t(u, v)$

Dans  $G$ , probabilité d'être sur  $v$  à l'instant  $t$  en partant de  $u$

# Marches Aléatoires

Liste ordonnées des verbes selon leur score de proxémie à partir de *déshabiller*  
( $t = 4$ )

1) dépouiller*	11) dévoiler	21) arracher	31) tailler	41) dire
2) défaire*	12) dégager	22) délacer*	32) développer	42) ouvrir
3) démunir*	13) étaler*	23) voler	33) déchirer	43) déployer
4) déshabiller*	14) ôter*	24) dénouer	34) déchausser	44) dépiauter
5) dévêtir*	15) enlever	25) desserrer	35) vider	45) déposséder
6) révéler	16) défrusquer	26) médire*	36) peler	46) frustrer
7) découvrir*	17) afficher	27) débarrasser	37) dépourvoir	47) indiquer
8) montrer*	18) désaffubler	28) exposer	38) défubler	48) démontrer
9) dénuder*	19) écorcher*	29) couper	39) priver	49) manifester
10) dégarnir*	20) démasquer*	30) exhiber*	40) trahir	50) détacher

**Table :** Les cinquante verbes qui ont la plus forte probabilité d'être atteints en 4 pas depuis *déshabiller* (sur 10000 verbes).

La Proxémie :  $P_G^t(u, v)$

Dans  $G$ , probabilité d'être sur  $v$  à l'instant  $t$  en partant de  $u$





$P_G^t(u, v)$  = Dans  $G$ , probabilité d'être sur  $v$  à l'instant  $t$  en partant de  $u$ .

- ▶ **Marches longues**: la probabilité d'atteindre un sommet  $v$  ne dépend que du degrés de  $v$ :

$$\lim_{t \rightarrow \infty} P_G^t(u, v) = \frac{\deg_G(v)}{\sum_{r \in V} \deg_G(r)} = \pi_G(v)$$

$P_G^t(u, v)$  = Dans  $G$ , probabilité d'être sur  $v$  à l'instant  $t$  en partant de  $u$ .

- ▶ **Marches longues:** la probabilité d'atteindre un sommet  $v$  ne dépend que du degrés de  $v$ :

$$\lim_{t \rightarrow \infty} P_G^t(u, v) = \frac{\deg_G(v)}{\sum_{r \in V} \deg_G(r)} = \pi_G(v)$$

- ▶ **Marches courtes:** forte probabilité de rester dans un cluster:

- $P_G^t(u, v) > \pi_G(v)$  si  $u$  et  $v$  sont dans un même cluster
- $P_G^t(u, v) < \pi_G(v)$  sinon

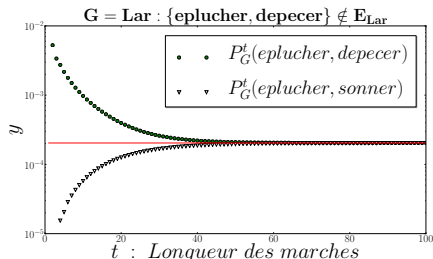
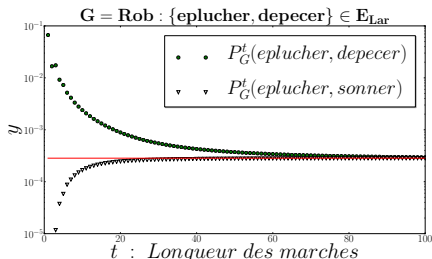
$P_G^t(u, v)$  = Dans  $G$ , probabilité d'être sur  $v$  à l'instant  $t$  en partant de  $u$ .

- **Marches longues:** la probabilité d'atteindre un sommet  $v$  ne dépend que du degrés de  $v$ :

$$\lim_{t \rightarrow \infty} P_G^t(u, v) = \frac{\deg_G(v)}{\sum_{r \in V} \deg_G(r)} = \pi_G(v)$$

- **Marches courtes:** forte probabilité de rester dans un cluster:

- $P_G^t(u, v) > \pi_G(v)$  si  $u$  et  $v$  sont dans un même cluster
- $P_G^t(u, v) < \pi_G(v)$  sinon



$P_G^t(u, v)$  = Dans  $G$ , probabilité d'être sur  $v$  à l'instant  $t$  en partant de  $u$ .

- **Marches longues:** la probabilité d'atteindre un sommet  $v$  ne dépend que du degrés de  $v$ :

$$\lim_{t \rightarrow \infty} P_G^t(u, v) = \frac{\deg_G(v)}{\sum_{r \in V} \deg_G(r)} = \pi_G(v)$$

- **Marches courtes:** forte probabilité de rester dans un cluster:

- $P_G^t(u, v) > \pi_G(v)$  si  $u$  et  $v$  sont dans un même cluster
- $P_G^t(u, v) < \pi_G(v)$  sinon

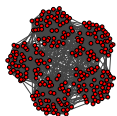
$$t = 5: \quad \text{CONF}_G(\mathbf{u}, \mathbf{v}) = \frac{P_G^5(\mathbf{u}, \mathbf{v})}{P_G^5(\mathbf{u}, \mathbf{v}) + \pi_G(\mathbf{v})}$$

$\text{CONF}_G$  est une mesure de proximité normalisée basée sur les marches aléatoires courtes dans  $G$ :

- Beaucoup de chemins courts entre  $u$  et  $v$  :  $\text{CONF}_G(u, v) > 0,5$ ;
- Peu de chemins courts entre  $u$  et  $v$  :  $\text{CONF}_G(u, v) < 0,5$ ;

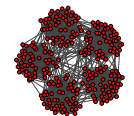
# Confluence: Expérience contrôlée

# Confluence: Expérience contrôlée

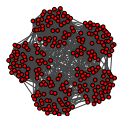


$$G_a = (V, E_a)$$

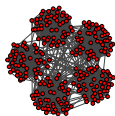
# Confluence: Expérience contrôlée



$$G_b = (V, E_b)$$

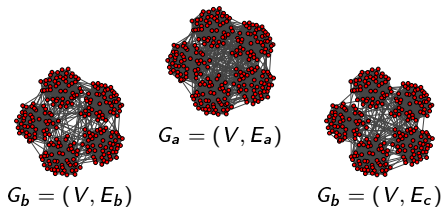


$$G_a = (V, E_a)$$



$$G_b = (V, E_c)$$

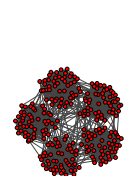
# Confluence: Expérience contrôlée



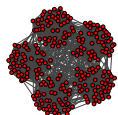
$$E_b \cap E_c = \emptyset \Rightarrow$$
$$GED(G_b, G_c) = 1$$



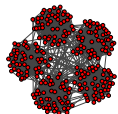
# Confluence: Expérience contrôlée



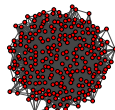
$G_b = (V, E_b)$



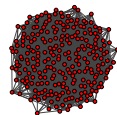
$G_a = (V, E_a)$



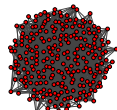
$G_b = (V, E_c)$



$G_b^R = (V, E_b^R)$



$G_a^R = (V, E_a^R)$

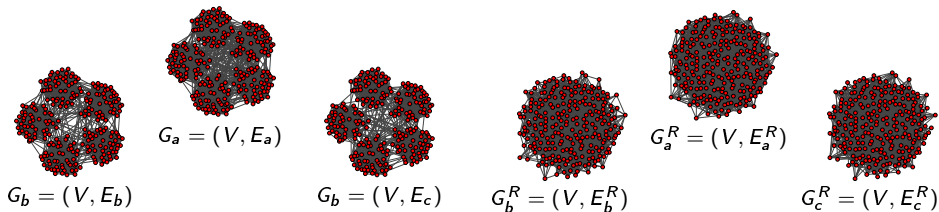


$G_c^R = (V, E_c^R)$

$$E_b \cap E_c = \emptyset \Rightarrow \\ GED(G_b, G_c) = 1$$

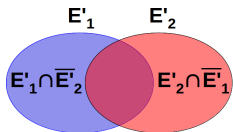
$$E_b^R \cap E_c^R = \emptyset \Rightarrow \\ GED(G_b^R, G_c^R) = 1$$

# Confluence: Expérience contrôlée

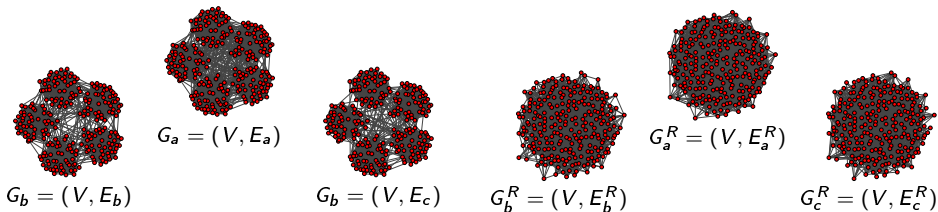


$$E_b \cap E_c = \emptyset \Rightarrow \\ GED(G_b, G_c) = 1$$

$$E_b^R \cap E_c^R = \emptyset \Rightarrow \\ GED(G_b^R, G_c^R) = 1$$

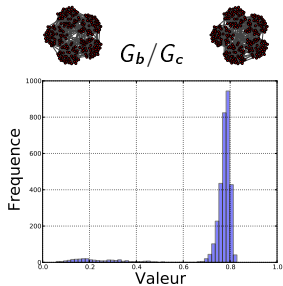
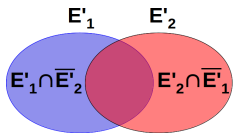


# Confluence: Expérience contrôlée

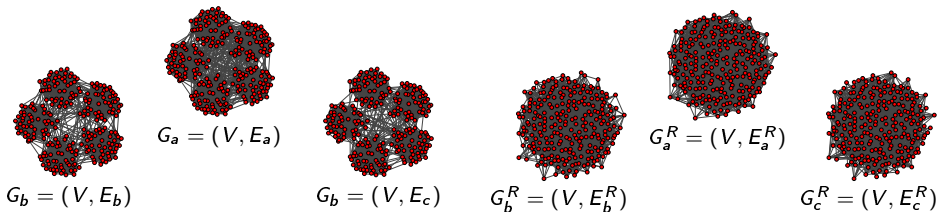


$E_b \cap E_c = \emptyset \Rightarrow$   
 $GED(G_b, G_c) = 1$

$E_b^R \cap E_c^R = \emptyset \Rightarrow$   
 $GED(G_b^R, G_c^R) = 1$

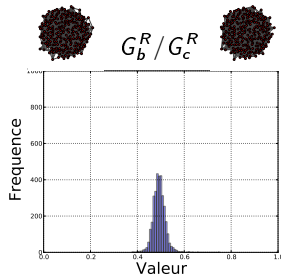
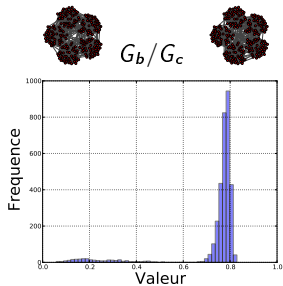
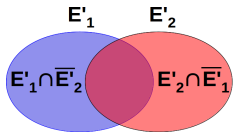


# Confluence: Expérience contrôlée



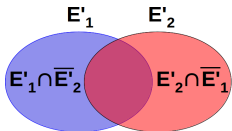
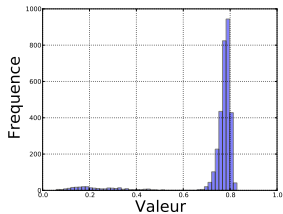
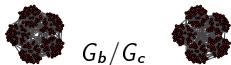
$E_b \cap E_c = \emptyset \Rightarrow$   
 $GED(G_b, G_c) = 1$

$E_b^R \cap E_c^R = \emptyset \Rightarrow$   
 $GED(G_b^R, G_c^R) = 1$

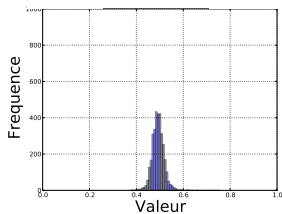
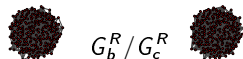


# Confluence: Expérience contrôlée

$$GED(G_b, G_c) = 1$$



$$GED(G_b^R, G_c^R) = 1$$

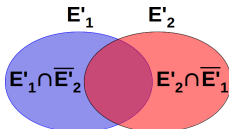
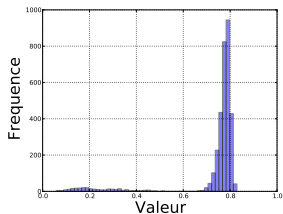
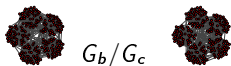


# Confluence: Expérience contrôlée

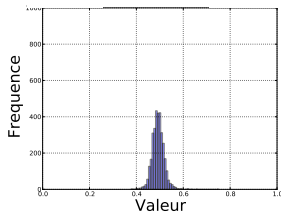
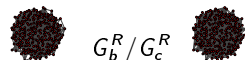
$$\mu(\mathbf{G}_1, \mathbf{G}_2) =$$

$$\frac{1}{|E_1 \cap \bar{E}_2| + |E_2 \cap \bar{E}_1|} \left( \sum_{\{u,v\} \in (E_2 \cap \bar{E}_1)} \text{CONF}_{G_1}(\{u,v\}) + \sum_{\{u,v\} \in (E_1 \cap \bar{E}_2)} \text{CONF}_{G_2}(\{u,v\}) \right)$$

$$GED(G_b, G_c) = 1$$



$$GED(G_b^R, G_c^R) = 1$$



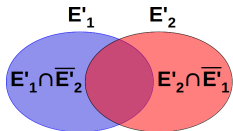
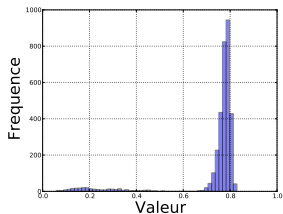
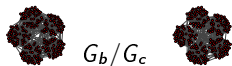
# Confluence: Expérience contrôlée

$$\mu(\mathbf{G}_1, \mathbf{G}_2) =$$

$$\frac{1}{|E_1 \cap \bar{E}_2| + |E_2 \cap \bar{E}_1|} \left( \sum_{\{u,v\} \in (E_2 \cap \bar{E}_1)} \text{CONF}_{G_1}(\{u,v\}) + \sum_{\{u,v\} \in (E_1 \cap \bar{E}_2)} \text{CONF}_{G_2}(\{u,v\}) \right)$$

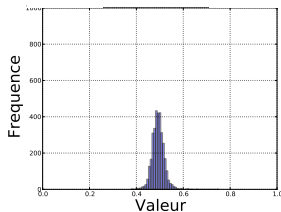
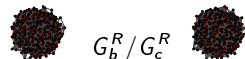
$$\mu(G_b, G_c) = 0.74$$

$$\text{GED}(G_b, G_c) = 1$$



$$\mu(G_b^R, G_c^R) = 0.49$$

$$\text{GED}(G_b^R, G_c^R) = 1$$



# Comparer les réseaux lexicaux



# Comparer les réseaux lexicaux

$$\mu(\mathbf{G}_1, \mathbf{G}_2) = \frac{1}{|E_1 \cap \bar{E}_2| + |E_2 \cap \bar{E}_1|} \left( \sum_{\{u,v\} \in (E_2 \cap \bar{E}_1)} \text{CONF}_{\mathbf{G}_1}(\{u,v\}) + \sum_{\{u,v\} \in (E_1 \cap \bar{E}_2)} \text{CONF}_{\mathbf{G}_2}(\{u,v\}) \right)$$

Robert / Larousse

# Comparer les réseaux lexicaux

$$\mu(\mathbf{G}_1, \mathbf{G}_2) = \frac{1}{|E_1 \cap \overline{E_2}| + |\overline{E_2} \cap E_1|} \left( \sum_{\{u,v\} \in (E_2 \cap \overline{E_1})} \text{CONF}_{G_1}(\{u,v\}) + \sum_{\{u,v\} \in (E_1 \cap \overline{E_2})} \text{CONF}_{G_2}(\{u,v\}) \right)$$

Robert / Larousse

$$\text{GED}(G'_{Rob}, G'_{Lar}) = 0.47$$



$$\text{GED}(G'^R_{Rob}, G'^R_{Lar}) = 0.47$$



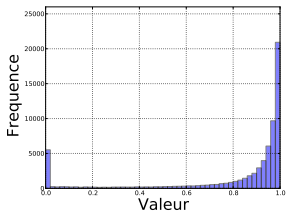
# Comparer les réseaux lexicaux

$$\mu(\mathbf{G}_1, \mathbf{G}_2) =$$

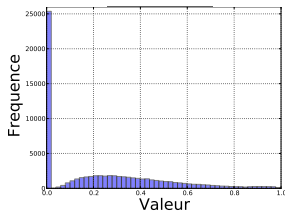
$$\frac{1}{|E_1 \cap \bar{E}_2| + |E_2 \cap \bar{E}_1|} \left( \sum_{\{u,v\} \in (E_2 \cap \bar{E}_1)} \text{CONF}_{G_1}(\{u,v\}) + \sum_{\{u,v\} \in (E_1 \cap \bar{E}_2)} \text{CONF}_{G_2}(\{u,v\}) \right)$$

Robert / Larousse

$$\text{GED}(G'_{Rob}, G'_{Lar}) = 0.47$$



$$\text{GED}(G'^R_{Rob}, G'^R_{Lar}) = 0.47$$



# Comparer les réseaux lexicaux

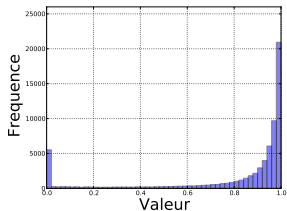
$$\mu(\mathbf{G}_1, \mathbf{G}_2) =$$

$$\frac{1}{|E_1 \cap \overline{E_2}| + |\overline{E_2} \cap E_1|} \left( \sum_{\{u,v\} \in (E_2 \cap \overline{E_1})} \text{CONF}_{G_1}(\{u,v\}) + \sum_{\{u,v\} \in (E_1 \cap \overline{E_2})} \text{CONF}_{G_2}(\{u,v\}) \right)$$

Robert / Larousse

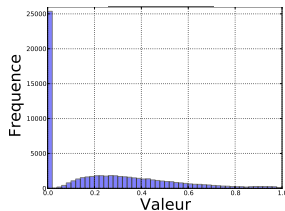
$$\mu(\mathbf{G}'_{\text{Rob}}, \mathbf{G}'_{\text{Lar}}) = 0.80$$

$$\text{GED}(\mathbf{G}'_{\text{Rob}}, \mathbf{G}'_{\text{Lar}}) = 0.47$$



$$\mu(\mathbf{G}'^{\text{R}}_{\text{Rob}}, \mathbf{G}'^{\text{R}}_{\text{Lar}}) = 0.24$$

$$\text{GED}(\mathbf{G}'^{\text{R}}_{\text{Rob}}, \mathbf{G}'^{\text{R}}_{\text{Lar}}) = 0.47$$



# Comparer les réseaux lexicaux

# Comparer les réseaux lexicaux

Table : Pédigrés de quelques graphes lexicaux

Graphes lexicaux		n	m	$\langle k \rangle$	C	$L_{lcc}$	$\lambda (r^2)$
Lar	Adjectifs	5510	21147	7,68	0,21	4,92	-2,06 (0,88)
	Noms	12159	31601	5,20	0,20	6,10	-2,39 (0,88)
	Verbes	5377	22042	8,20	0,17	4,61	-1,94 (0,88)
Rob	Adjectifs	7693	20011	5,20	0,14	5,26	-2,05 (0,94)
	Noms	24570	55418	4,51	0,11	6,08	-2,34 (0,94)
	Verbes	7357	26567	7,22	0,12	4,59	-2,01 (0,93)
Jdm	Adjectifs	9859	30087	6,10	0,16	5,44	-2,24 (0,90)
	Noms	29213	56381	3,86	0,14	6,48	-2,66 (0,93)
	Verbes	7658	22260	5,81	0,14	5,06	-2,08 (0,89)
Wik	Adjectifs	6960	6594	1,89	0,15	8,48	-2,46 (0,95)
	Noms	43206	37661	1,74	0,13	10,56	-2,51 (0,89)
	Verbes	7203	7497	2,08	0,25	9,22	-2,28 (0,92)

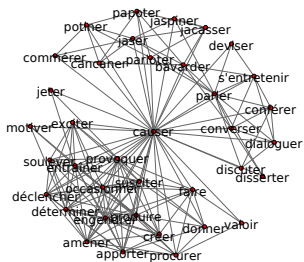
# Comparer les réseaux lexicaux

Table : Mesurer la similarité structurelle entre réseaux lexicaux

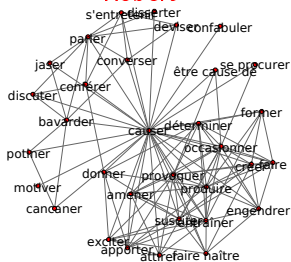
GED ( $\mu$ ) ( $\mu^R$ ) sur les paires de graphes			
$G_1/G_2$	$Rob_A$	$Jdm_A$	$Wik_A$
$Lar_A$	<b>0,45</b> ( <b>0,76</b> ) (0,34)	<b>0,47</b> ( <b>0,71</b> ) (0,38)	<b>0,75</b> ( <b>0,41</b> ) (0,06)
$Rob_A$		<b>0,51</b> ( <b>0,70</b> ) (0,29)	<b>0,71</b> ( <b>0,42</b> ) (0,05)
$Jdm_A$			<b>0,54</b> ( <b>0,43</b> ) (0,03)
$G_1/G_2$	$Rob_N$	$Jdm_N$	$Wik_N$
$Lar_N$	<b>0,48</b> ( <b>0,70</b> ) (0,20)	<b>0,48</b> ( <b>0,70</b> ) (0,19)	<b>0,72</b> ( <b>0,31</b> ) (0,03)
$Rob_N$		<b>0,47</b> ( <b>0,70</b> ) (0,13)	<b>0,71</b> ( <b>0,29</b> ) (0,02)
$Jdm_N$			<b>0,46</b> ( <b>0,32</b> ) (0,01)
$G_1/G_2$	$Rob_V$	$Jdm_V$	$Wik_V$
$Lar_V$	<b>0,48</b> ( <b>0,73</b> ) (0,40)	<b>0,46</b> ( <b>0,70</b> ) (0,39)	<b>0,78</b> ( <b>0,25</b> ) (0,05)
$Rob_V$		<b>0,47</b> ( <b>0,70</b> ) (0,37)	<b>0,78</b> ( <b>0,25</b> ) (0,06)
$Jdm_V$			<b>0,55</b> ( <b>0,31</b> ) (0,04)

# Comparer les réseaux lexicaux

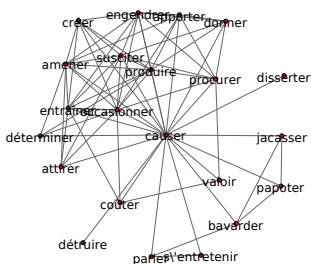
## Larousse



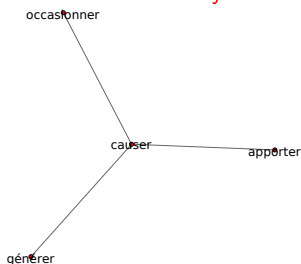
## Robert



## Jeux de mots



## Wiktionary







**Hypothèse** : La synonymie reflète la structure sémantique du lexique commune aux membres d'une même communauté linguistique

**Constat** : Il existe de nombreux désaccords entre membres d'une même communauté linguistique:

$$\text{GED}(G_1, G_2) \approx 0.50$$

- **Beaucoup** de paires de mots sont synonymes pour l'individu  $I_1$  et pas pour l'individu  $I_2$



**Hypothèse** : La synonymie reflète la structure sémantique du lexique commune aux membres d'une même communauté linguistique

**Incompatible ?**

**Constat** : Il existe de nombreux désaccords entre membres d'une même communauté linguistique:

**GED( $G_1, G_2$ )  $\approx 0.50$**

- **Beaucoup** de paires de mots sont synonymes pour l'individu  $I_1$  et pas pour l'individu  $I_2$



**Hypothèse** : La synonymie reflète la structure sémantique du lexique commune aux membres d'une même communauté linguistique

**Incompatible ?**

**Constat** : Il existe de nombreux désaccords entre membres d'une même communauté linguistique:

**$GED(G_1, G_2) \approx 0.50$**

- **Beaucoup** de paires de mots sont synonymes pour l'individu  $I_1$  et pas pour l'individu  $I_2$



**Non**

**Parceque** : Ces désaccords sont des désaccords faibles :

**$\mu(G_1, G_2) \approx 0.70$  (accord au niveau structurel)**

**Hypothèse** : La synonymie reflète la structure sémantique du lexique commune aux membres d'une même communauté linguistique

**Ce n'est pas la somme du sens de ses arêtes qui donne le sens d'une relation lexicale, mais le sens de la relation lexicale dans la globalité de sa structure qui donne du sens à ses arêtes.**

**Hypothèse** : La synonymie reflète la structure sémantique du lexique commune aux membres d'une même communauté linguistique

**Ce n'est pas la somme du sens de ses arêtes qui donne le sens d'une relation lexicale, mais le sens de la relation lexicale dans la globalité de sa structure qui donne du sens à ses arêtes.**

**L'analyse d'une relation lexicale doit être faite au niveau de la structure globale dessinée par la relation. C'est ce que la mesure  $\mu$  peut faire: avec un niveau de représentation adéquat (la confluence), elle réconcilie les jugements portés par deux juges différents sur une même relation lexicale.**

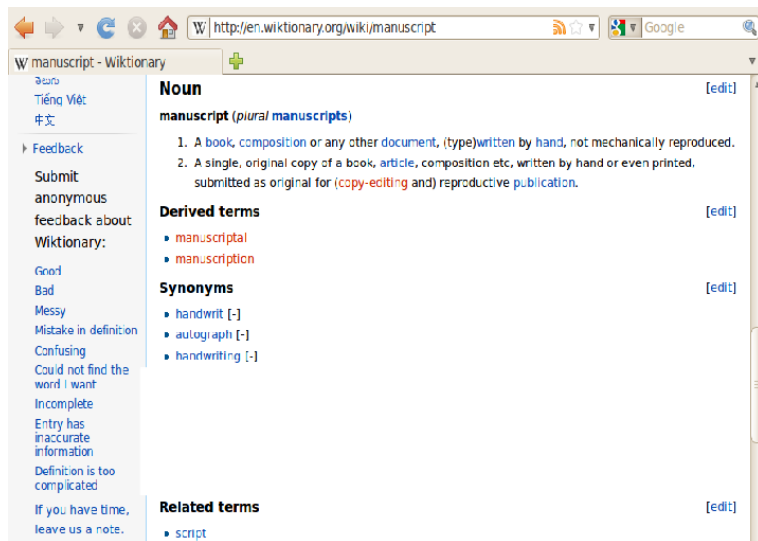
- 1 Structure des réseaux lexicaux
- 2 Mesurer la similarité structurelle entre réseaux lexicaux

## Applications

- 3 WISIGOTH : Enrichissement de lexiques approvisionnés par les foules**
- 4 SLAM : Substitution lexicale en contexte
- 5 CELLO : Détection de communautés pour la Recherche d'Information
- 6 SKILLEX : Un score d'habileté lexicale

- ▶ Sajous F., Navarro E., Gaume B., Prévot L., Chudy Y (2011). WISIGOTH Semi-Automatic Enrichment of Crowdsourced Synonymy Networks: an Application to Wiktionary **in** proc. of LRE Language Resources and Evaluation : Special Issue on Collaboratively Constructed Language Resources
- ▶ Sajous F., Navarro E., Gaume B. (2011). Enrichissement de lexiques sémantiques approvisionnés par les foules : le système WISIGOTH appliqué à Wiktionary **in** Journal TAL, 52-1 (11-35).

# Enrichissement de lexiques



The screenshot shows a web browser window displaying the Wiktionary page for the word "manuscript". The browser's address bar shows the URL "http://en.wiktionary.org/wiki/manuscript". The page content is as follows:

**Noun** [edit]

**manuscript** (*plural manuscripts*)

1. A **book**, **composition** or any other **document**, (type)written by **hand**, not mechanically reproduced.
2. A single, original copy of a book, **article**, composition etc, written by hand or even printed, submitted as original for (**copy-editing** and) reproductive **publication**.

**Derived terms** [edit]

- **manuscriptal**
- **manuscription**

**Synonyms** [edit]

- **handwrit** [-]
- **autograph** [-]
- **handwriting** [-]

**Related terms** [edit]

- **script**

On the left side of the page, there is a sidebar with various options and feedback links, including "Feedback", "Submit anonymous", "feedback about", "Wiktionary:", "Good", "Bad", "Messy", "Mistake in definition", "Confusing", "Could not find the word I want", "Incomplete", "Entry has inaccurate information", "Definition is too complicated", and "If you have time, leave us a note."



# Enrichissement de lexiques

The screenshot shows the English Wiktionary page for the word "manuscript". The browser address bar displays "http://en.wiktionary.org/wiki/manuscript". The page content includes:

- Noun** [edit]
- manuscript** (plural **manuscripts**)
- 1. A **book**, **composition** or any other **document**, (type)written by **hand**, not mechanically reproduced.
- 2. A single, original copy of a book, **article**, composition etc, written by hand or even printed, submitted as original for (**copy-editing** and) reproductive **publication**.
- Derived terms** [edit]:
  - **manuscriptal**
  - **manuscription**
- Synonyms** [edit]:
  - **handwrit** [-]
  - **autograph** [-]
  - **handwriting** [-]
- Wisigoth: candidate synonyms for manuscript (N)** [ More synonyms ]
- Table of candidate synonyms:

<i>script</i> [ + ] [ x ]	<i>copy</i> [ + ] [ x ]	<i>journal</i> [ + ] [ x ]	<i>register</i> [ + ] [ x ]
<i>writing</i> [ + ] [ x ]	<i>scenario</i> [ + ] [ x ]	<i>screenplay</i> [ + ] [ x ]	<i>inscription</i> [ + ] [ x ]
<i>signature</i> [ + ] [ x ]	<i>original</i> [ + ] [ x ]		
- Propose your own synonym here:
- Related terms** [edit]:
  - **script**

The left sidebar contains navigation options for different languages (Tiếng Việt, 中文) and feedback options (Submit anonymous feedback about Wiktionary: Good, Bad, Messy, Mistake in definition, Confusing, Could not find the word I want, Incomplete, Entry has inaccurate information, Definition is too complicated, If you have time, leave us a note).

- 1 Structure des réseaux lexicaux
- 2 Mesurer la similarité structurelle entre réseaux lexicaux

## Applications

- 3 WISIGOTH : Enrichissement de lexiques approvisionnés par les foules
- 4 **SLAM : Substitution lexicale en contexte**
- 5 CELLO : Détection de communautés pour la Recherche d'Information
- 6 SKILLEX : Un score d'habileté lexicale

- ▶ Desalle., Gaume B., Duvignau, K. (2009). SLAM : Solution Lexicale Automatique de Métaphores *in* revue : Traitement Automatique des Langues
- ▶ Desalle Y., Navarro E., Chudy Y., Magistry P., Gaume B. (2014). BACANAL : Balades Aléatoires Courtes pour ANALyses Lexicales Application à la substitution lexicale *in* Proc. of TALN-20 2014 : Atelier SEMDIS, Marseille, France.

Le but de la tâche est la substitution lexicale : remplacer un mot  $\omega$  dans une phrase  $P$  par un autre mot tout en préservant au maximum le sens de la phrase  $P$ .

**Exemple :** *Et cette confiance **fonde** la responsabilité du praticien.*

**Gold Semdis14 :** créer, forger, constituer, justifier, être à la base, entraîner, assurer, impliquer, baser, instaurer, induire, définir, être à l'origine, établir, installer, poser, supporter

**Exemple :** *Et cette confiance **fonde** la responsabilité du praticien.*

**Gold Semdis14 :** créer, forger, constituer, justifier, être à la base, entraîner, assurer, impliquer, baser, instaurer, induire, définir, être à l'origine, établir, installer, poser, supporter

Dans **Gjdm<sub>v</sub>**, **fonder** a 15 voisins : affermir, appuyer, asseoir, assoir, baser, bâtir, constituer, créer, former, instaurer, instituer, justifier, édifier, élever, ériger

**Exemple :** *Et cette confiance **fonde** la responsabilité du praticien.*

**Gold Semdis14** : créer, forger, constituer, justifier, être à la base, entraîner, assurer, impliquer, baser, instaurer, induire, définir, être à l'origine, établir, installer, poser, supporter

Dans **Gjdm<sub>v</sub>**, **fonder** a 15 voisins : affermir, appuyer, asseoir, assoir, baser, bâtir, **constituer**, **créer**, former, **instaurer**, instituer, **justifier**, édifier, élever, ériger

**Exemple :** *Et cette confiance **fonde** la responsabilité du praticien.*

**Gold Semdis14 :** créer, forger, constituer, justifier, être à la base, entraîner, assurer, impliquer, baser, instaurer, induire, définir, être à l'origine, établir, installer, poser, supporter

Dans **Gjdm<sub>v</sub>**, **fonder** a 15 voisins : affermir, appuyer, asseoir, assoir, baser, bâtir, **constituer**, **créer**, former, **instaurer**, instituer, **justifier**, édifier, élever, ériger

Dans **Glm10<sub>v</sub>**, **fonder** a 583 voisins: NC.espoir.Dep.obj (freq=144, IM=120.347), NC.société.Dep.obj (freq=138, IM=59.6531), NC.famille.Dep.obj (freq=98, IM=77.4457), NC.revue.Dep.obj (freq=85, IM=317.996), V.venir.Gov.Prep/de (freq=82, IM=7.29996), NC.principe.Dep.suj (freq=82, IM=84.03), NC.compagnie.Dep.obj (freq=79, IM=120.454), NC.parti.Dep.obj (freq=78, IM=46.1816), NC.association.Dep.obj (freq=78, IM=100.757), NC.valeur.Dep.suj (freq=76, IM=73.5628)

**Exemple :** *Et cette confiance fonde la responsabilité du praticien.*

**Gold Semdis14 :** créer, forger, constituer, justifier, être à la base, entraîner, assurer, impliquer, baser, instaurer, induire, définir, être à l'origine, établir, installer, poser, supporter

Dans **Gjdm<sub>v</sub>** : **CONF<sub>Gjdm<sub>v</sub></sub>(fonder, u)**

Dans **Glm10<sub>v</sub>** : **CONF<sub>Glm10<sub>v</sub></sub>(NC.responsabilite.Dep.obj, u)**



**Exemple :** *Et cette confiance **fonde** la responsabilité du praticien.*

**Gold Semdis14 :** créer, forger, constituer, justifier, être à la base, entraîner, assurer, impliquer, baser, instaurer, induire, définir, être à l'origine, établir, installer, poser, supporter

Dans **Gjdm<sub>v</sub>** : **CONF<sub>Gjdm<sub>v</sub></sub>(fonder, u)**

Dans **Glm10<sub>v</sub>** : **CONF<sub>Glm10<sub>v</sub></sub>(NC.responsabilite.Dep.obj, u)**

**établir, constituer, créer, former, instituer, assurer, mettre, instaurer, poser, construire**

**Exemple :** *Et cette confiance **fonde** la responsabilité du praticien.*

**Gold Semdis14 :** créer, forger, constituer, justifier, être à la base, entraîner, assurer, impliquer, baser, instaurer, induire, définir, être à l'origine, établir, installer, poser, supporter

Dans **Gjdm<sub>V</sub>** : **CONF<sub>Gjdm<sub>V</sub></sub>(fonder, u)**

Dans **Glm10<sub>V</sub>** : **CONF<sub>Glm10<sub>V</sub></sub>(NC.responsabilite.Dep.obj, u)**

**établir, constituer, créer, former, instituer, assurer, mettre, instaurer, poser, construire**

Meilleure méthode parmi celles qui ont concouru à SEMDIS-2014 (sur un jeu d'évaluation de 300 phrases)

**Exemple :** *Et cette confiance **fonde** la responsabilité du praticien.*

**Gold Semdis14** : créer, forger, constituer, justifier, être à la base, entraîner, assurer, impliquer, baser, instaurer, induire, définir, être à l'origine, établir, installer, poser, supporter

Dans **Gjdm<sub>V</sub>** : **CONF<sub>Gjdm<sub>V</sub></sub>(fonder, u)**

Dans **Glm10<sub>V</sub>** : **CONF<sub>Glm10<sub>V</sub></sub>(NC.responsabilite.Dep.obj, u)**

**établir, constituer, créer, former, instituer, assurer, mettre, instaurer, poser, construire**

Meilleure méthode parmi celles qui ont concouru à SEMDIS-2014 (sur un jeu d'évaluation de 300 phrases)

**Démétaphorisation** : Il **déshabille** la pomme → peler, éplucher, ...

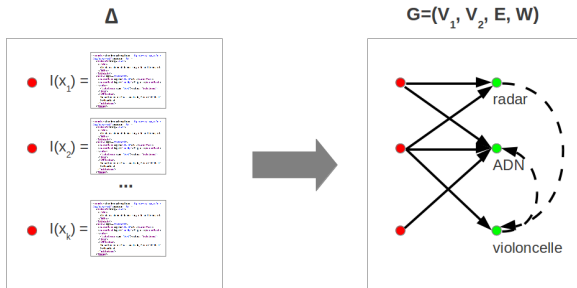
- 1 Structure des réseaux lexicaux
- 2 Mesurer la similarité structurelle entre réseaux lexicaux

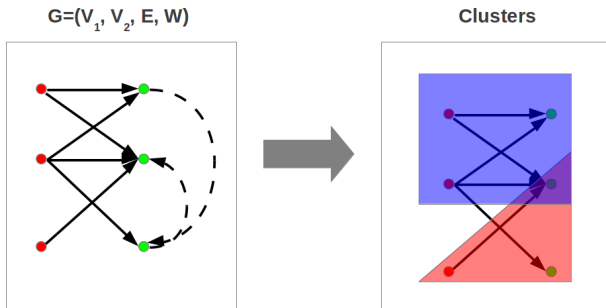
## Applications

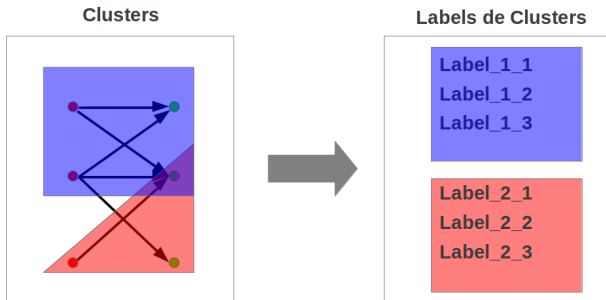
- 3 WISIGOTH : Enrichissement de lexiques approvisionnés par les foules
- 4 SLAM : Substitution lexicale en contexte
- 5 CELLO : Détection de communautés pour la Recherche d'Information**
- 6 SKILLEX : Un score d'habileté lexicale

► Navarro, E., Chudy, Y., Gaume, B., Cabanac, G., and Pinel-Sauvagnat, K.(2011). Kodex ou comment organiser les résultats d'une recherche d'information par détection de communautés sur un graphe biparti ? In CORIA'11, Avignon, pages 25–40. ARIA.











News  
Kodex over some newspapers

euro

search

Searching  
Guardian\_search

Clustering  
Infomap

Germany, World news, Angela Merkel, Europe, Eurozone crisis, European Union, Economics

Labour, UK news, Politics, Ed Balls, Labour conference 2013, Gordon Brown, Damian McBride

Sport, Football, Premier League, Fulham, Chelsea, Liverpool, Luis Suárez

- 1 [Greece pleads for more time to cut public sector jobs - live](#) ●  
September 24, 2013: Greece pleads for more time to cut public sector - live  
<http://www.theguardian.com/business/2013/sep/24/german-coalition-haggling-begins-as-greek-state-workers-strike-again-live>
- 2 [If Europe's future is German, Greece has only one hope for salvation | Costas Douzinas](#) ●  
September 24, 2013: If Europe's future is German, Greece has only one hope for salvation  
<http://www.theguardian.com/commentisfree/2013/sep/24/german-angela-merkel-austerity>
- 3 [The Spanish language school that delivers lessons on a budget](#) ○  
September 24, 2013: The Spanish language school that delivers lessons on a budget  
<http://www.theguardian.com/education/2013/sep/24/spanish-state-funded-language-schools>
- 4 [Labour conference - Ed Balls's speech: Politics live blog](#) ●  
September 23, 2013: Labour conference - Ed Balls's speech: Politics live blog  
<http://www.theguardian.com/politics/2013/sep/23/labour-conference-ed-balls-speech-politics-live-blog>
- 5 [Markets fall as Angela Merkel faces 'difficult' coalition talks - as it happened](#) ●  
September 23, 2013: Markets fall as Angela Merkel begins 'difficult' coalition talks - as it happened  
<http://www.theguardian.com/business/2013/sep/23/markets-calm-after-merkels-election-triumph-live>
- 6 [Want to be a successful world leader? Do it the Angela Merkel way](#) ●  
September 23, 2013: Want to be a successful world leader? Do it the Angela Merkel way  
<http://www.theguardian.com/world/shortcuts/2013/sep/23/angela-merkel-successful-world-leader-germany>
- 7 [Ed Balls's message to Labour: we will be less useless next time - probably](#) ●  
September 23, 2013: Ed Balls's message to Labour: we will be less useless next time - probably

**Gain de CELLO versus Terrier R:35% P:27% F:22%**

News  
Kodex over some newspapers

euro

search

Searching  
Guardian\_search

Clustering  
Infomap

Germany, World news, Angela Merkel, Europe, Eurozone crisis, European Union, Economics

Labour, UK news, Politics, Ed Balls, Labour conference 2013, Gordon Brown, Damian McBride

Sport, Football, Premier League, Fulham, Chelsea, Liverpool, Luis Suárez

- [1 Greece pleads for more time to cut public sector jobs - live](#) ●  
September 24, 2013: Greece pleads for more time to cut public sector - live  
<http://www.theguardian.com/business/2013/sep/24/german-coalition-haggling-begins-as-greek-state-workers-strike-again-live>
- [2 If Europe's future is German, Greece has only one hope for salvation | Costas Douzinas](#) ●  
September 24, 2013: If Europe's future is German, Greece has only one hope for salvation  
<http://www.theguardian.com/commentisfree/2013/sep/24/german-angela-merkel-austerity>
- [3 The Spanish language school that delivers lessons on a budget](#) ○  
September 24, 2013: The Spanish language school that delivers lessons on a budget  
<http://www.theguardian.com/education/2013/sep/24/spanish-state-funded-language-schools>
- [4 Labour conference - Ed Balls's speech: Politics live blog](#) ●  
September 23, 2013: Labour conference - Ed Balls's speech: Politics live blog  
<http://www.theguardian.com/politics/2013/sep/23/labour-conference-ed-balls-speech-politics-live-blog>
- [5 Markets fall as Angela Merkel faces 'difficult' coalition talks - as it happened](#) ●  
September 23, 2013: Markets fall as Angela Merkel begins 'difficult' coalition talks - as it happened  
<http://www.theguardian.com/business/2013/sep/23/markets-calm-after-merkels-election-triumph-live>
- [6 Want to be a successful world leader? Do it the Angela Merkel way](#) ●  
September 23, 2013: Want to be a successful world leader? Do it the Angela Merkel way  
<http://www.theguardian.com/world/shortcuts/2013/sep/23/angela-merkel-successful-world-leader-germany>
- [7 Ed Balls's message to Labour: we will be less useless next time - probably](#) ●  
September 23, 2013: Ed Balls's message to Labour: we will be less useless next time - probably

# SKILLEX : Un score d'habileté lexicale

- 1 Structure des réseaux lexicaux
- 2 Mesurer la similarité structurelle entre réseaux lexicaux

## Applications

- 3 WISIGOTH : Enrichissement de lexiques approvisionnés par les foules
- 4 SLAM : Substitution lexicale en contexte
- 5 CELLO : Détection de communautés pour la Recherche d'Information
- 6 SKILLEX : Un score d'habileté lexicale**

► Desalle, Y , Gaume B, Duvignau K., Hintat, C., Shu- Kai, H., Magistry, P., Nespoulous, J.-L. (2014). Skillex, an action labelling efficiency score: the case for French and Mandarin in Proceedings of COGSCI14, THE ANNUAL MEETING OF THE COGNITIVE SCIENCE SOCIETY, Quebec.

# SKILLEX : Un score d'habileté lexicale

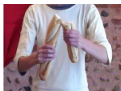
Protocole **APPROX** Gaume B., Duvignau K. (2004)



1 Démonteur des  
légos à la main



2 Couper du pain  
avec un couteau



3 Rompre du pain à  
la main



4 Déchirer un  
journal à la main



5 Scier un planche  
avec une scie



6 Hacher du persil  
avec un hachoir



7 Briser un verre  
avec un marteau



8 Déshabiller une  
poupée à la main



9 Explorer un  
ballon à la main



10 Ecorcer une  
buche à la main



11 Ecraser une  
tomate à la main



12 Emitter le pain  
dur à la main



13 Eplucher une  
banane à la main



14 Peler une carotte  
avec un éplucheur



15 Peler une  
orange à la main



16 Froisser une feuille  
de papier à la main

# SKILLEX : Un score d'habileté lexicale

Protocole **APPROX** Gaume B., Duvignau K. (2004)



1 Démont  
er des  
légos à  
la main



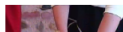
2 Couper  
du pain  
avec un  
couteau



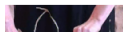
3 Rompre  
du pain à  
la main



4 Déchirer  
un  
journal à  
la main



9 Exploder  
un  
ballon à  
la main



10 Ecorcer  
une  
buche à  
la main



11 Ecraser  
une  
tomate à  
la main



12 Emitter  
du pain  
dur à la  
main



13 Eplucher  
une  
banane à  
la main



14 Peler  
une  
carotte  
avec un  
éplucheur



15 Peler  
une  
orange à  
la main

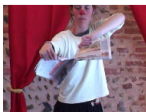


16 Froisser  
une  
feuille  
de papier  
à la main

Qu'est-ce qu'elle a fait la dame?  
Ce qu'elle a fait la dame, di[te]s le moi d'une autre manière, avec  
d'autres mots.

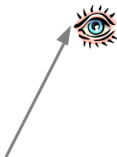
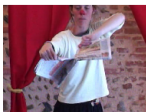
# SKILLEX : Un score d'habileté lexicale

## Actions



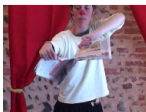
# SKILLEX : Un score d'habileté lexicale

Actions



# SKILLEX : Un score d'habileté lexicale

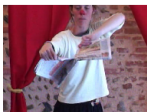
Actions



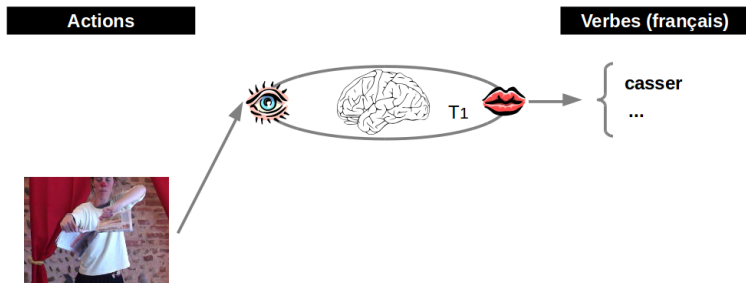


# SKILLEX : Un score d'habileté lexicale

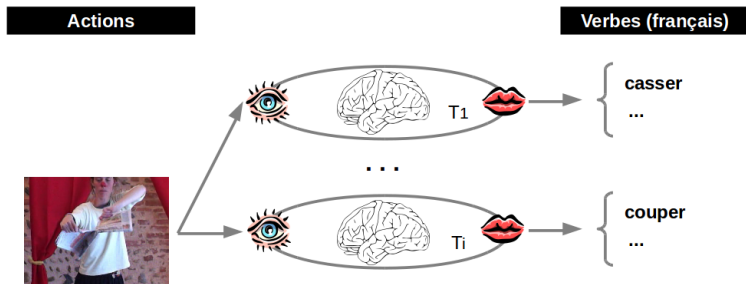
Actions



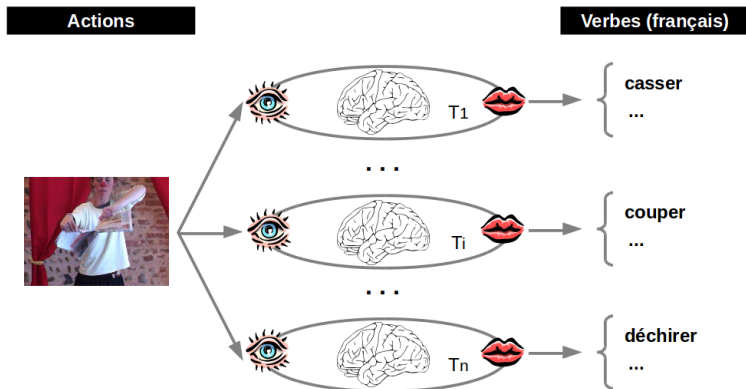
# SKILLEX : Un score d'habileté lexicale



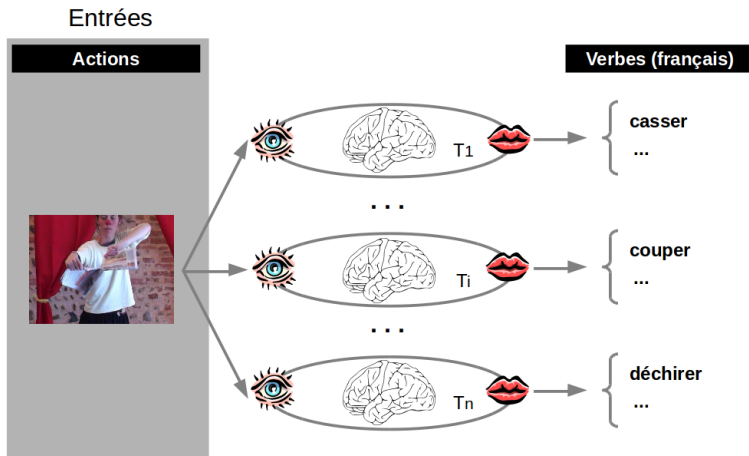
# SKILLEX : Un score d'habileté lexicale



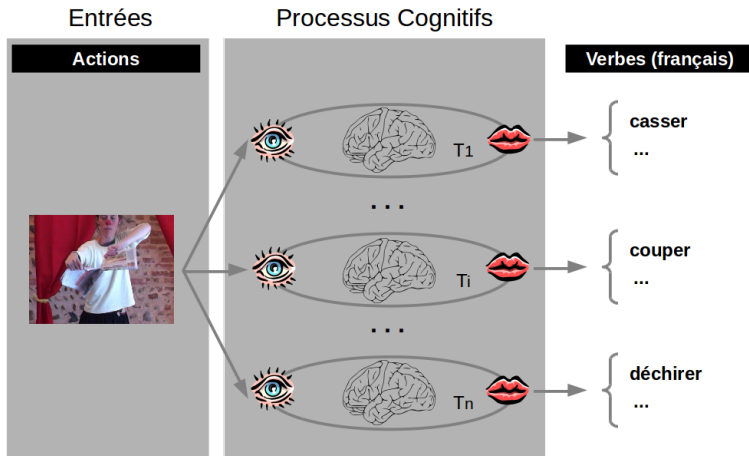
# SKILLEX : Un score d'habileté lexicale



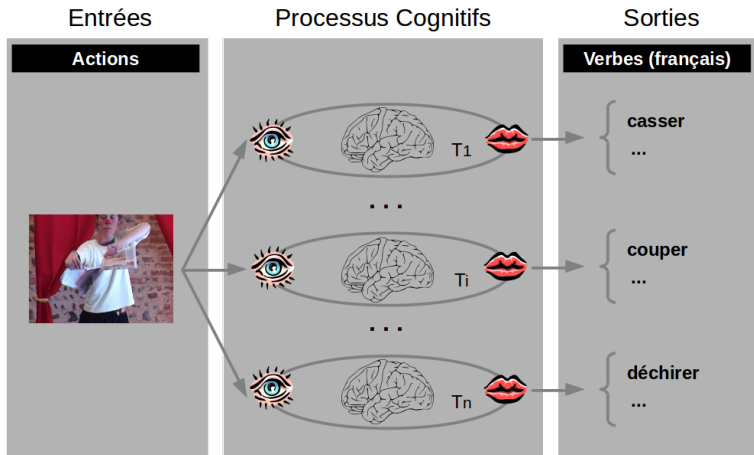
# SKILLEX : Un score d'habileté lexicale



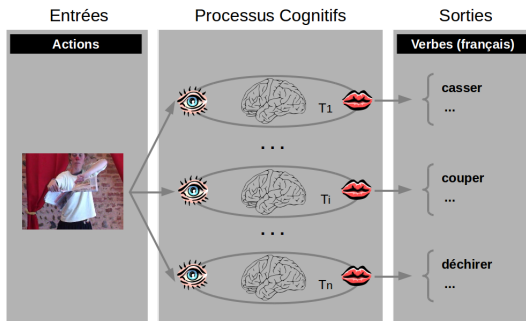
# SKILLEX : Un score d'habileté lexicale



# SKILLEX : Un score d'habileté lexicale

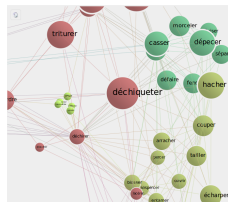
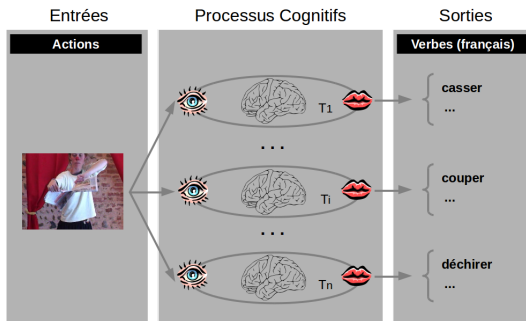


# SKILLEX : Un score d'habileté lexicale



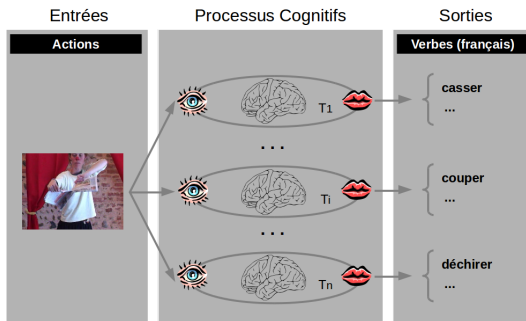


# SKILLEX : Un score d'habileté lexicale

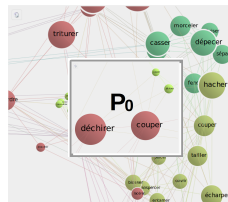




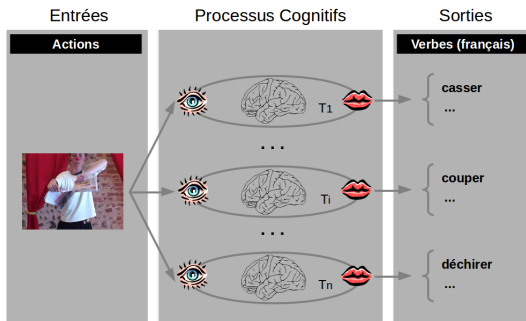
# SKILLEX : Un score d'habileté lexicale



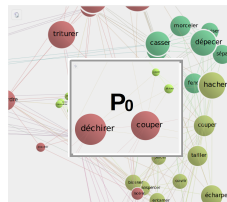
$$S_i = \frac{CONF_G(P_0, v_i)}{deg(v_i)}$$



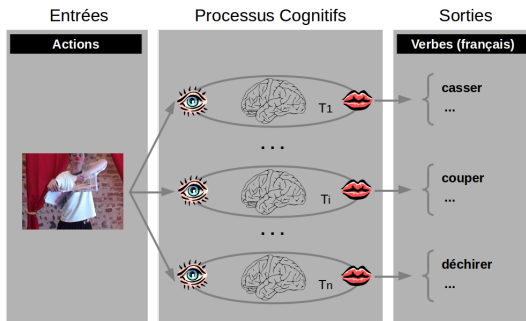
# SKILLEX : Un score d'habileté lexicale



$$S_i = \frac{CONF_G(P_0, v_i)}{deg(v_i)} \quad \left. \vphantom{\frac{CONF_G(P_0, v_i)}{deg(v_i)}} \right\} \text{Skillex}(i)$$

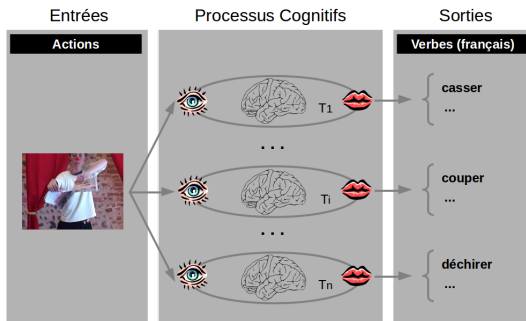


# SKILLEX : Un score d'habileté lexicale



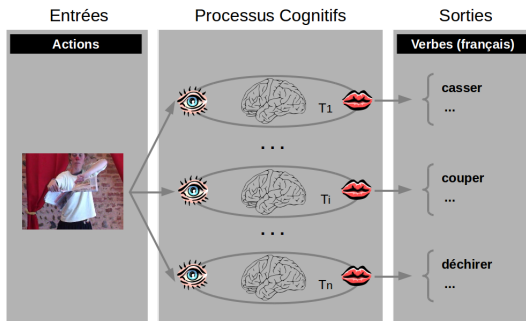
**Skillex(i)**

# SKILLEX : Un score d'habileté lexicale

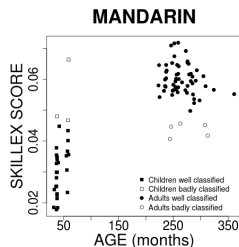


Skillex(i)

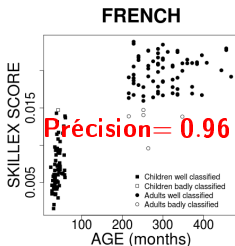
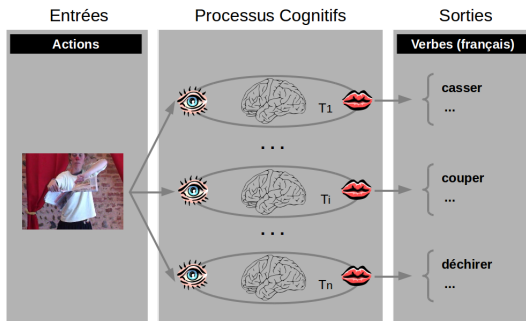
# SKILLEX : Un score d'habileté lexicale



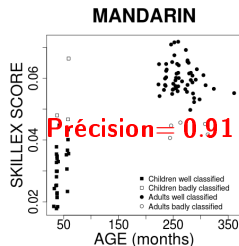
Skillex(i)



# SKILLEX : Un score d'habileté lexicale

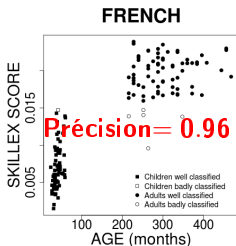
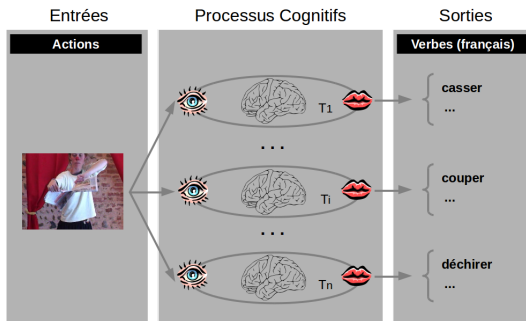


Skillex(i)



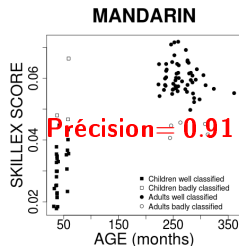


# SKILLEX : Un score d'habileté lexicale

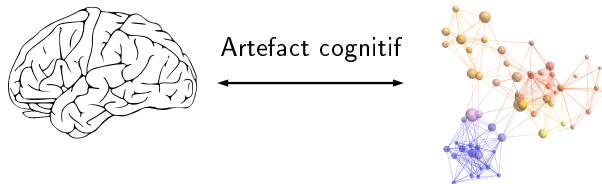
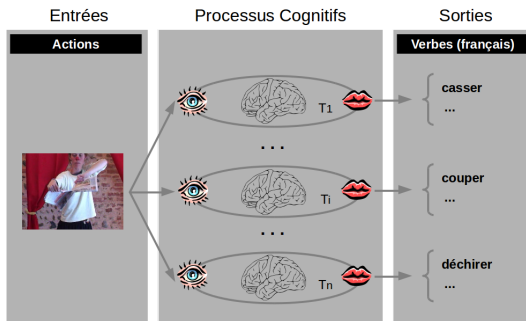


**Robuste**

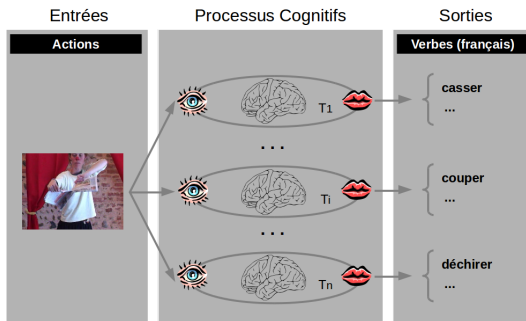
**Skillex(i)**



# SKILLEX : Un score d'habileté lexicale

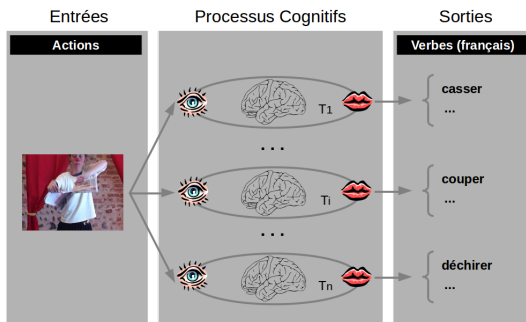


# SKILLEX : Un score d'habileté lexicale



Sujets Enfant/Sujets Adultes

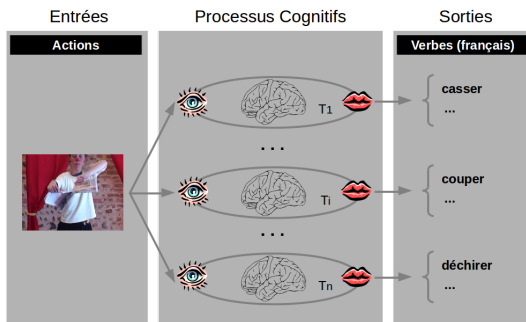
# SKILLEX : Un score d'habileté lexicale



Sujets Enfant/Sujets Adultes

Sujets Alzheimer/Sujets sans troubles

# SKILLEX : Un score d'habileté lexicale



Sujets Enfant/Sujets Adultes

Sujets Alzheimer/Sujets sans troubles

Sujets Enfants Asperger/Sujets Enfants sans troubles

- 1 Structure des réseaux lexicaux
- 2 Mesurer la similarité structurelle entre réseaux lexicaux

## Applications

- 3 WISIGOTH : Enrichissement de lexiques approvisionnés par les foules
- 4 SLAM : Substitution lexicale en contexte
- 5 CELLO : Détection de communautés pour la Recherche d'Information
- 6 SKILLEX : Un score d'habileté lexicale

