

DIC 9410

# Présentation du projet de recherche

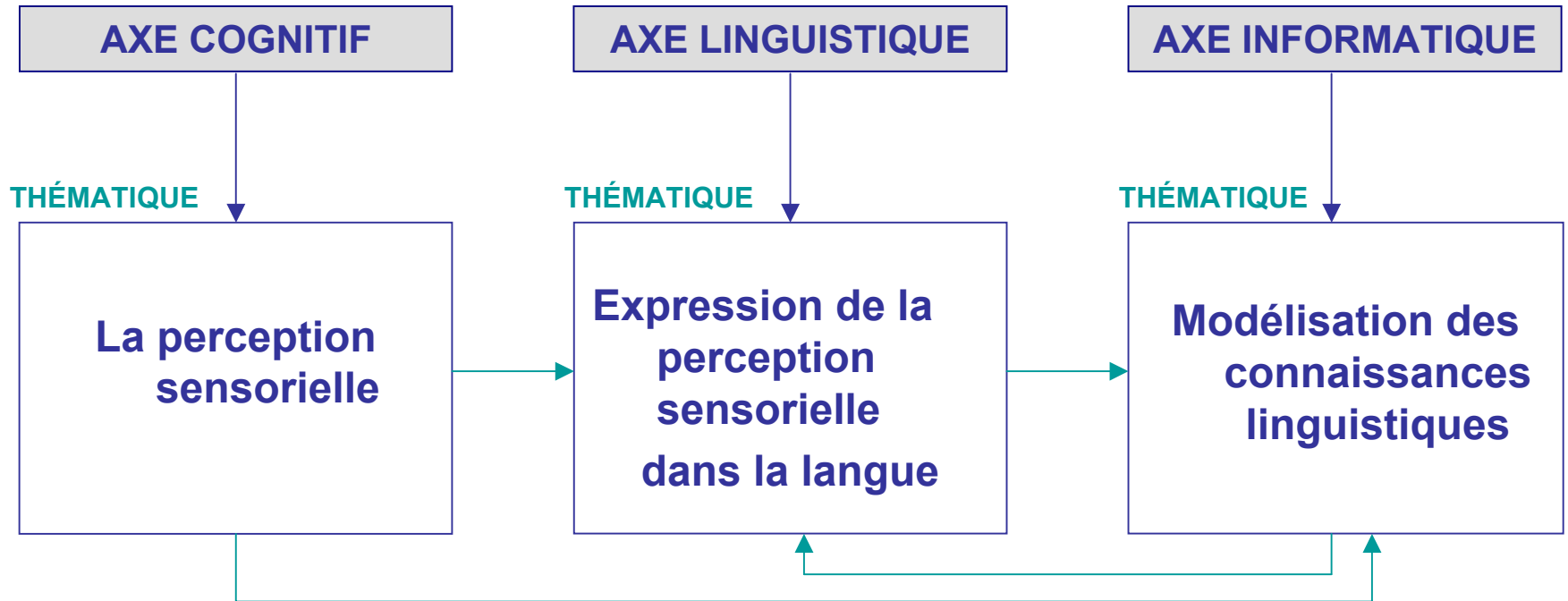
Vers un traitement automatique des  
langues : les cas des verbes de  
perception sensorielle en français

Sophie Piron

- Problématique
  - Axes et thématiques de la recherche
  - Objet d'étude
- Hypothèses
- Exposition des aspects de la recherche
  - Axe linguistique
  - Axe cognitif
  - Axe informatique

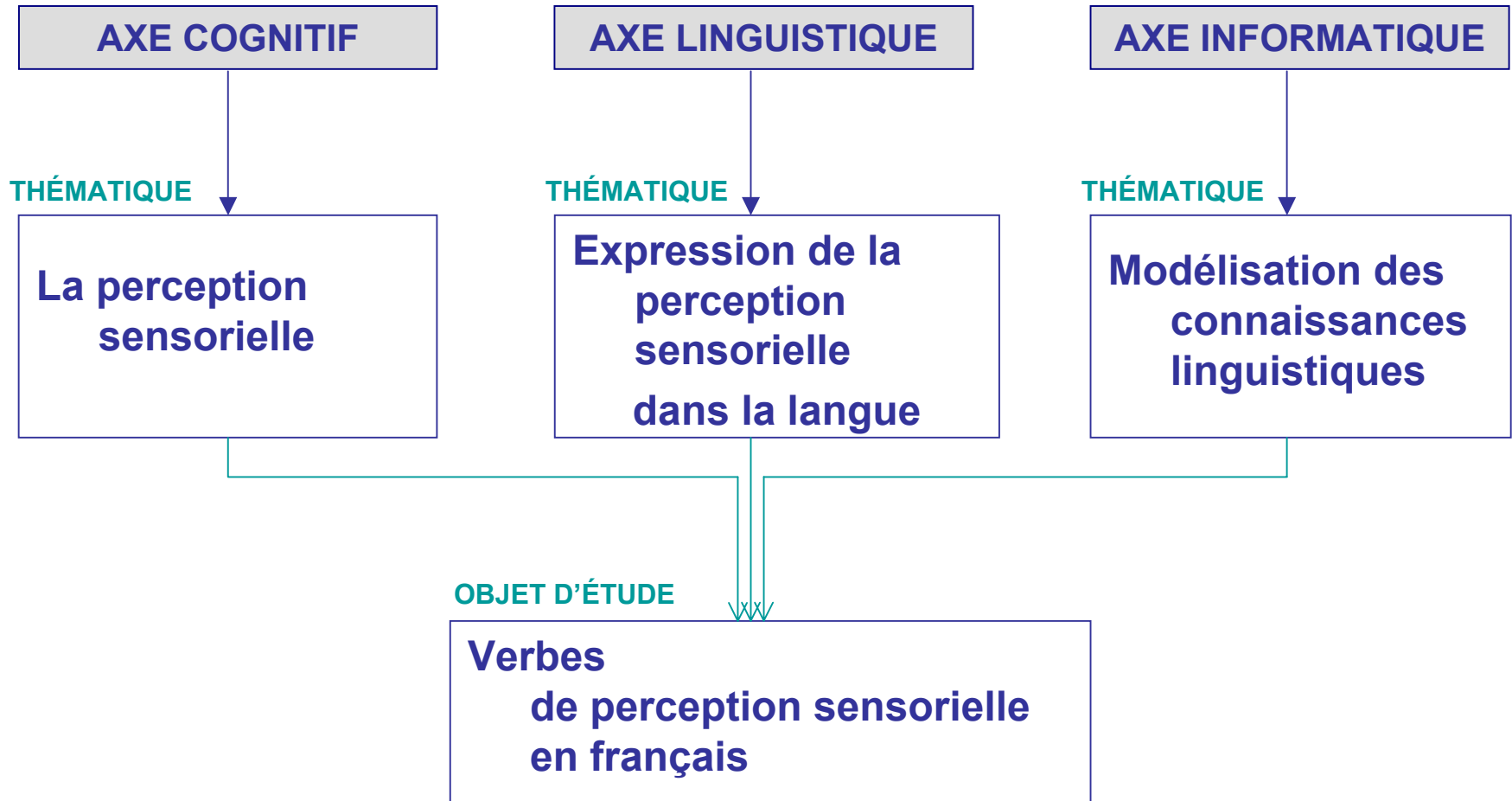
## Axes et thématiques de la recherche

---



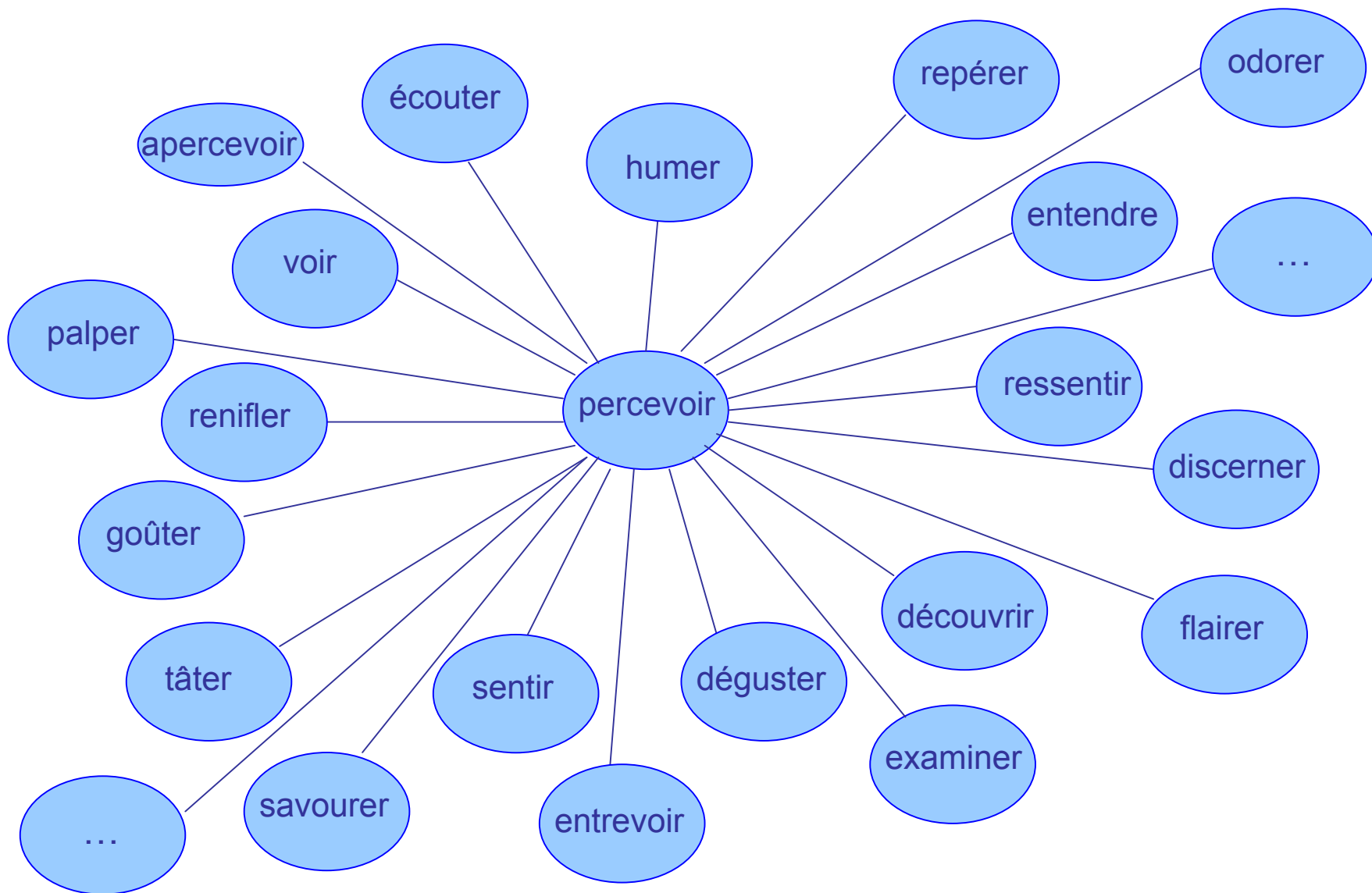
## Axes et thématiques de la recherche

---



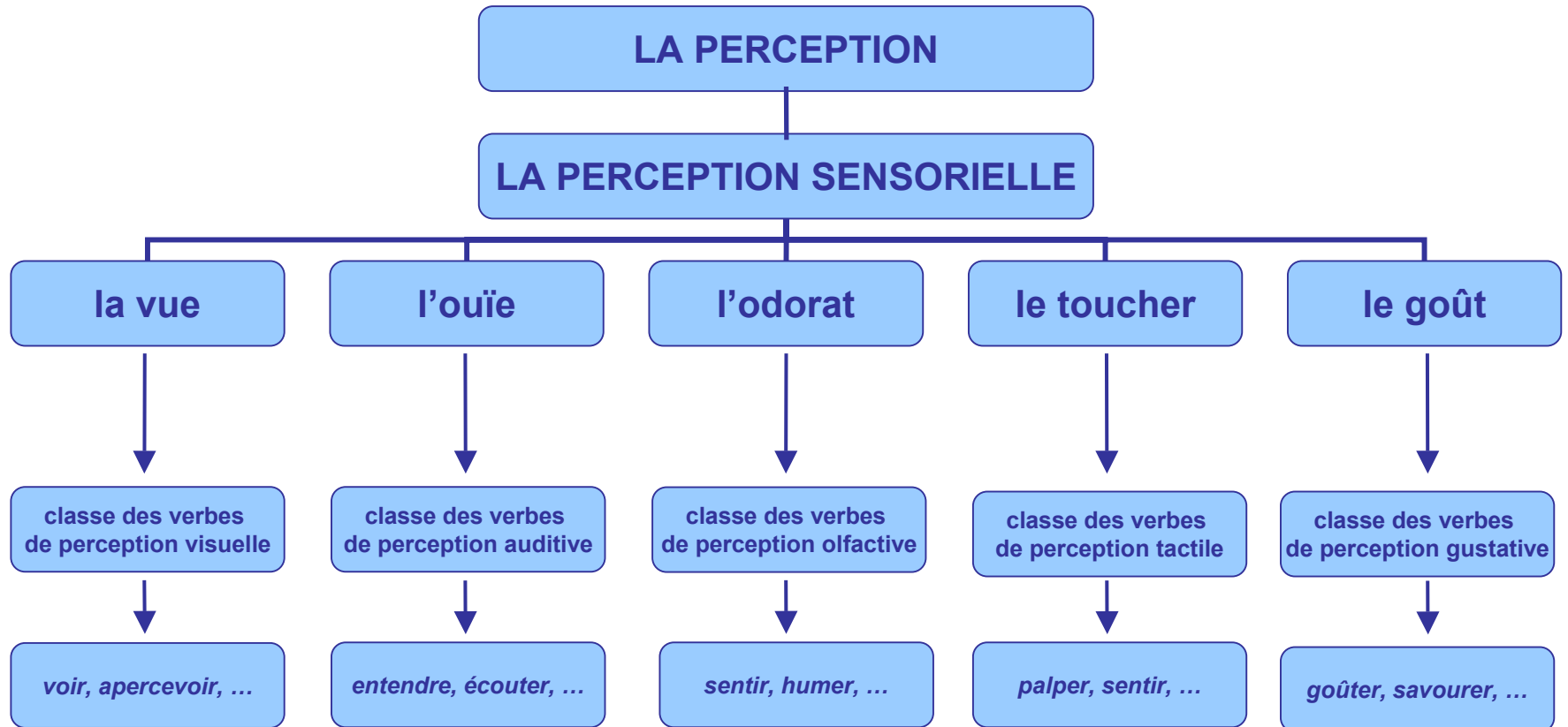
## Objet d'étude : les verbes de perception

---

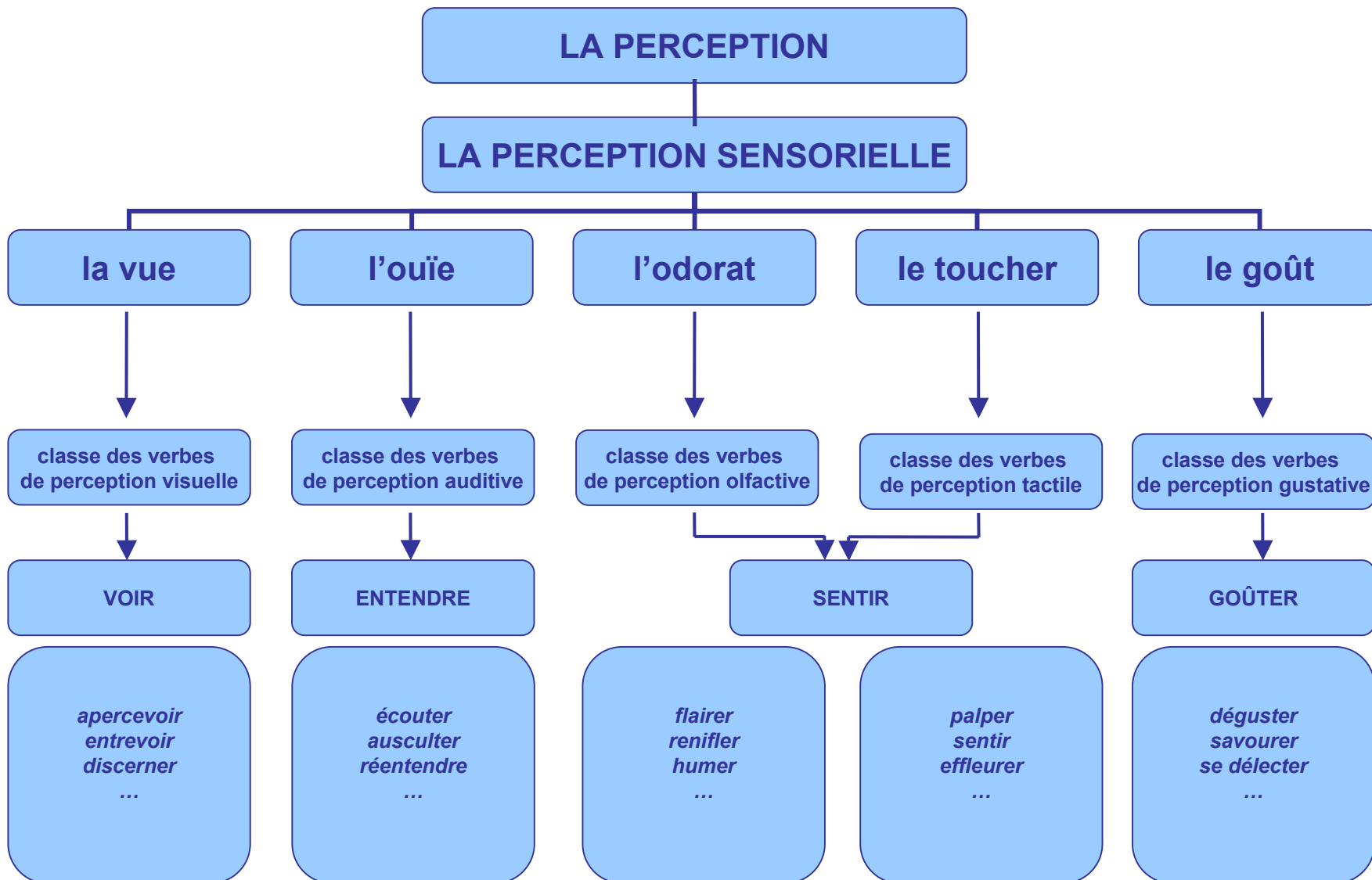


# Objet d'étude : les verbes de perception

---



# Objet d'étude : les verbes de perception



- **HYPOTHÈSE 1 – axes cognitif et linguistique**

L'expression de la perception sensorielle est directement influencée par la cognition sensorielle et les aspects physiologiques de celle-ci.

- **HYPOTHÈSE 2 – axes cognitif et linguistique**

L'analyse linguistique de la classe des verbes de perception sensorielle est plus précise grâce à l'introduction de réflexions sur la perception sensorielle.

- **HYPOTHÈSE 3 – axes cognitif, linguistique et informatique**

La modélisation des connaissances linguistiques (lexique computationnel, à savoir la composante lexicale d'un système de traitement automatique des langues) nécessite une réflexion linguistique approfondie et un ancrage en sémantique cognitive.



# Développement de la recherche

## Domaine : sciences cognitives

AXE COGNITIF

AXE LINGUISTIQUE

PSYCHOLOGIE EXPÉR.

LINGUISTIQUE

- fondements neuro-physiologiques
- fondements neuronaux
- cognition sensorielle

fondements historiques

champs sémantiques

étude de corpus

analyse linguistique

## Domaine : informatique

AXE INFORMATIQUE

FORMALISATION

APPLICATION

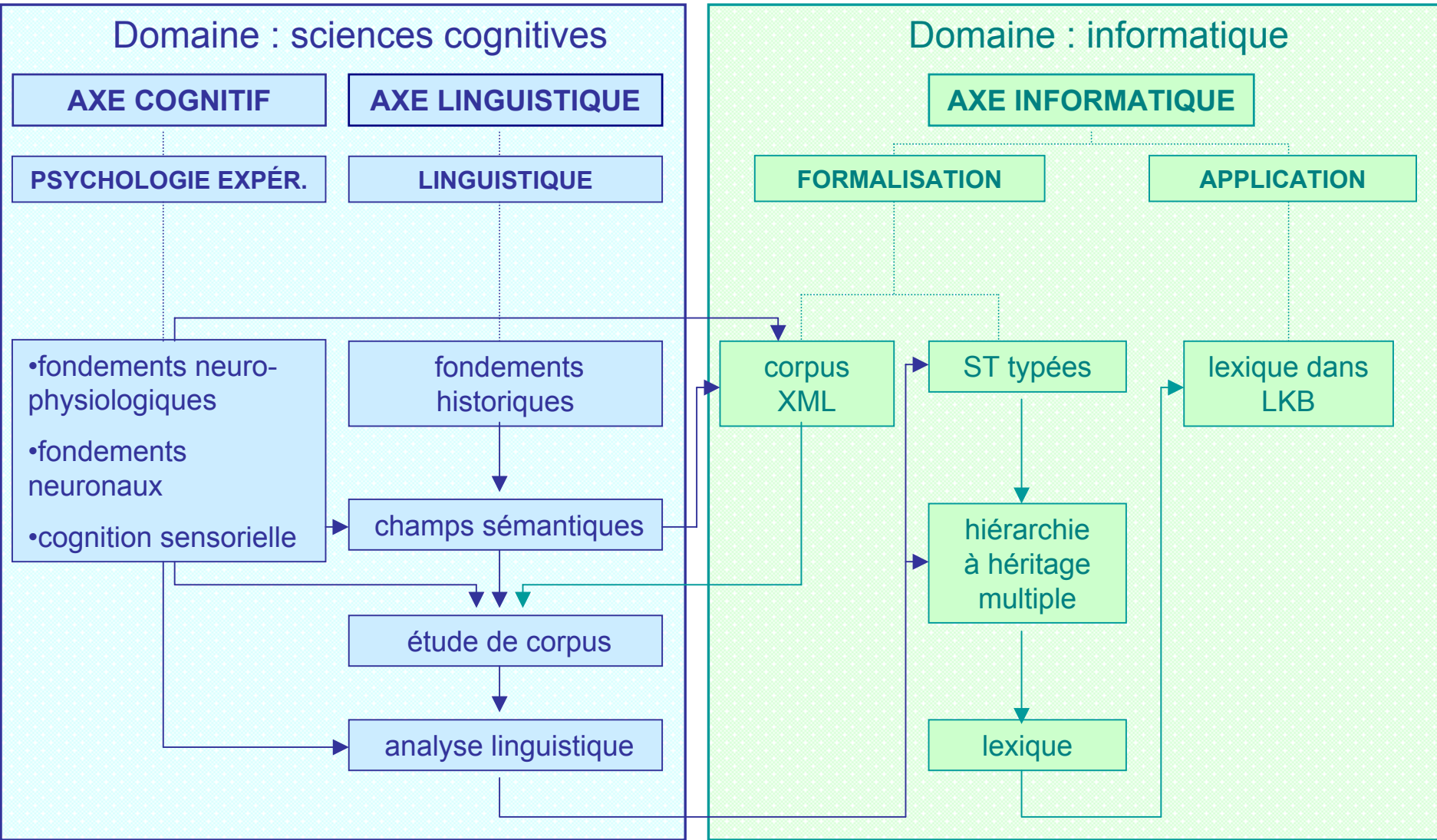
corpus XML

ST typées

lexique dans LKB

hiérarchie à héritage multiple

lexique



# Développement de la recherche

## Domaine : sciences cognitives

**AXE COGNITIF**

**AXE LINGUISTIQUE**

**PSYCHOLOGIE EXPÉR.**

**LINGUISTIQUE**

- fondements neuro-physiologiques
- fondements neuronaux
- cognition sensorielle

**fondements historiques**

**champs sémantiques**

## Domaine : informatique

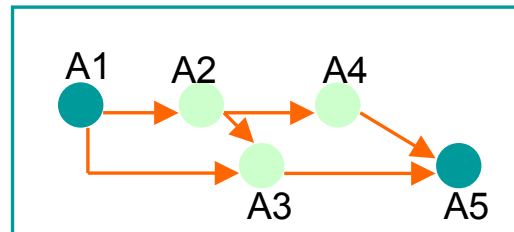
**AXE INFORMATIQUE**

**FORMALISATION**

**APPLICATION**

## Problématique

- Perte d'une partie de la compréhension du phénomène étudié
- Liste d'acceptations : liens parfois difficiles à établir

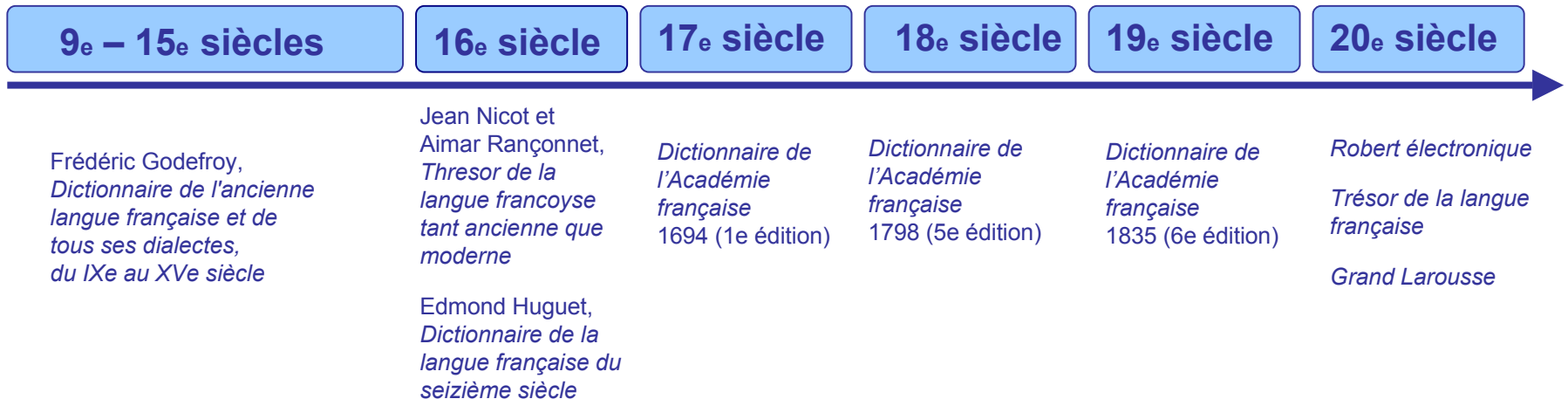


## Objectifs

- Faire apparaître les liens entre les acceptations
- Faire apparaître que l'expression d'un phénomène n'est jamais fixée
- Ébaucher la carte conceptuelle des verbes étudiés

### Méthodologie

- Consultation d'ouvrages lexicographiques de référence
- Regroupement des notices et des acceptions proposées



# Axe linguistique – fondements historiques – exemple du verbe *entendre*

## DOMAINE PHYSIQUE

tendre

## DOMAINE SENSORIEL

percevoir (ouïe)

apprendre par la rumeur

écouter

écouter un témoignage

exaucer

## DOMAINE MENTAL

attendre

vouloir

être attentif à qqn

s'occuper de qqch.

consentir

avoir une entente secrète

s'associer

sympathiser

comprendre

être habile dans qqch.

9<sup>e</sup> – 15<sup>e</sup> siècles

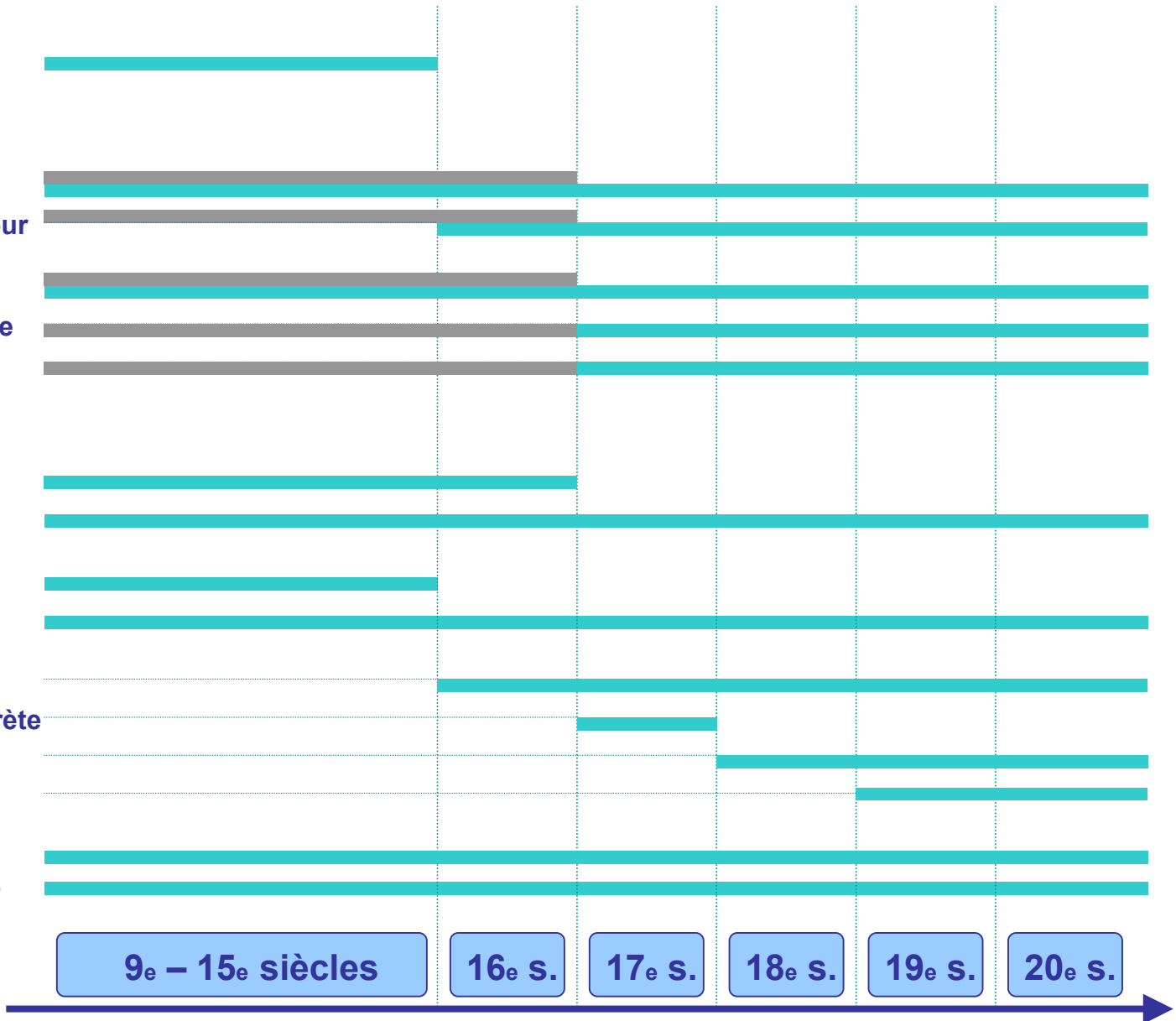
16<sup>e</sup> s.

17<sup>e</sup> s.

18<sup>e</sup> s.

19<sup>e</sup> s.

20<sup>e</sup> s.



# Développement de la recherche

## Domaine : sciences cognitives

AXE COGNITIF

AXE LINGUISTIQUE

PSYCHOLOGIE EXPÉR.

LINGUISTIQUE

•fondements neuro-physiologiques

•fondements neuronaux

•cognition sensorielle

fondements historiques

champs sémantiques

## Domaine : informatique

AXE INFORMATIQUE

FORMALISATION

APPLICATION

corpus XML

## Problématique

- Base empirique pour l'analyse linguistique  
(extraction de ressources)

## Objectifs

- Liste plus précise des emplois des verbes de perception sensorielle
- Informations sur les caractéristiques des emplois (fréquences relatives, restrictions sélectionnelles, etc.)
- Constitution de ressources linguistiques classables et consultables (mouvance de TEI P4, EAGLES, projet TIGER)

### DESCRIPTION

- Corpus d'échantillons
- Monolingue
- Synchronique
- Langue écrite
- Domaine général
- Registres littéraire et courant (journalistique)

**Citation** ✕

Les \*mâchoires étaient si serrées que, de la porte, Jacques \*entendit  
crisser les dents.  
MARTIN DU GARD, les Thibault, t. IV, p. 153.

Recherche de citations AVEC mâchoires ET AVEC entendit dans le texte

Précédent    Suivant    1 / 1    Copier

**STEP 1**    **STEP 2**    **STEP 3**

SELECT ONE LANGUAGE:    SELECT ONE/SEVERAL NEWSPAPERS:    CONCORDANCE SORT OPTION:

ENGLISH  
 FRENCH  
 PORTUGUESE  
 ITALIAN  
 GERMAN  
 SPANISH  
 DUTCH  
 NORWEGIAN  
 GREEK  
 SLOVAK  
 RUSSIAN

La Dernière Heure / Les Sports (Belgium)  
La Libre Belgique (Belgium)  
La Meuse (Belgium)  
Le Soir (Belgium)  
Cyberpresse (Canada)  
Dernières Nouvelles d'Alsace (France)  
L'Est Républicain (France)  
L'Humanité (France)  
La Voix du Nord (France)  
Le Bien Public (France)  
Le Figaro (France)  
Le Monde (France)  
Le Parisien (France)  
Le Progrès de Lyon (France)  
Le Télégramme (France)

Center,Right

TYPE YOUR QUERY HERE:  
<entendre>

search

[\[Info about Advanced Queries\]](#)

GlossaNet is developed and hosted by [CENTAL](#)



### CORPUS BRUT

- s les faits, l'auditorat réagit... lentement. {S} Un exemple: Olivier Pirson est entendu une première fois quatre jours après le décès de ses enfants. {S} On le s  
(*La Meuse*, 1er décembre 2001)
- avant dans le corps de son époux." {S} Une soeur aînée de l'accusé a aussi été entendue .{S} Dépressive, souvent hospitalisée, elle a été fréquemment tenue à l'  
(*La Meuse*, 1er décembre 2001)
- ssé un portait exemplaire de Christine, médecin urgentiste avec laquelle elle s'entendait très bien. {S} De son frère, elle a dit: "Il s'est senti moralement ob  
(*La Meuse*, 1er décembre 2001)
- Les mâchoires étaient si serrées que, de la porte, Jacques entendit crisser les dents.  
MARTIN DU GARD, *les Thibault*, t. IV, p. 153.
- e le rappelle Vincent Sclamender, " est causé par la chute de la pluie (ça sous-entend une certaine douceur dans les couches d'air supérieures) sur un sol gelé  
(*La Meuse*, 3 février 2003)
- [...] il me fit entendre que Mme Dupin trouvait mes visites trop fréquentes et me priait de les discontinuer.  
(Rousseau, *les Confessions*, VII)
- - [...] je n'entends pas que vous fassiez de dépense [...] je vous demande de ne me traiter qu'en ami.  
(Molière, *Monsieur de Pourceaugnac*, I, 7)

## Schéma XSD

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
  elementFormDefault="qualified">
  <xs:element name="entendre">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="section" maxOccurs="unbounded"/>
      </xs:sequence>
    </xs:complexType>
  <xs:element name="ex" type="xs:string"/>
  <xs:element name="sem-verbe" type="xs:string"/>
  <xs:element name="pattern-syn" type="xs:string"/>
  <xs:element name="suppl-syn-sem" type="xs:string"/>
  <xs:element name="section">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="ex"/>
        <xs:element ref="sem-verbe"/>
        <xs:element ref="pattern-syn"/>
        <xs:element ref="suppl-syn-sem"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
  ...
</xs:schema>
```

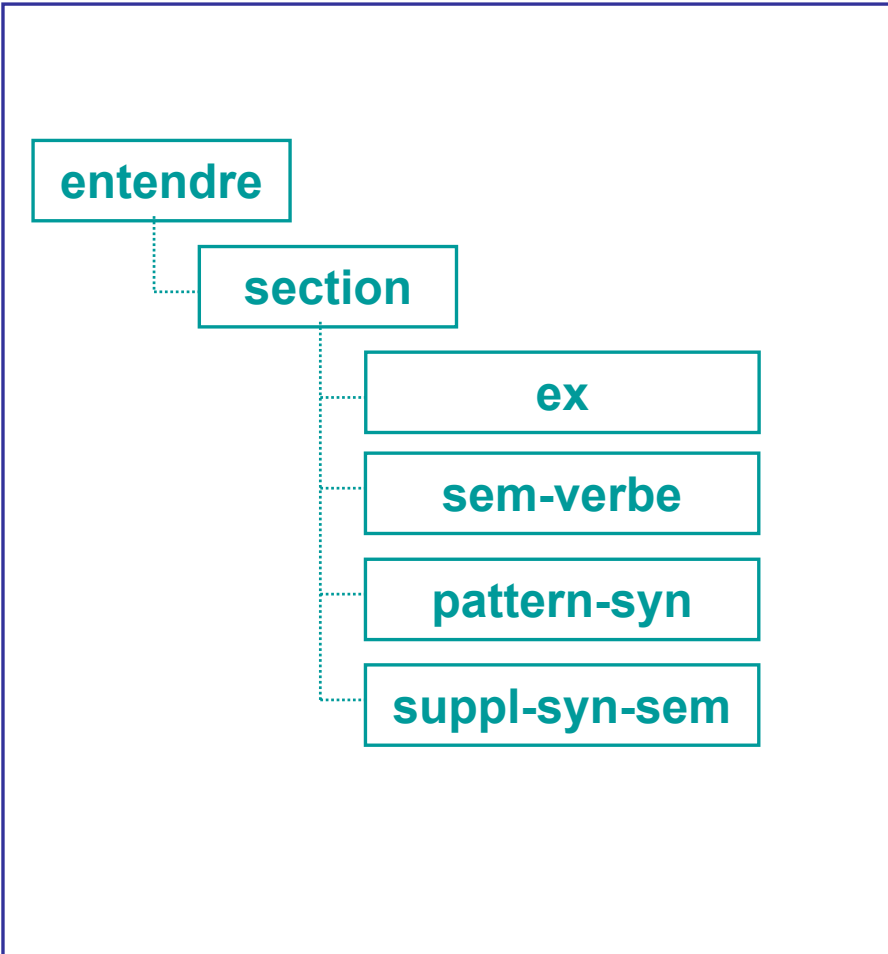
## Document XML

```
<entendre xmlns:xsi="http://www.w3.org/2001/XMLSchema-
instance"
  xsi:noNamespaceSchemaLocation="D:\Documents\xml\ente
ndre_stat.xsd">
  ...
  <section id="entendre_GrandRobert_cit20_291">
    <ex>Dans la lingerie, les femmes repassaient;
    [on entendait le bruit des fers] (...) des coups sourds
    sur les molletons (...)
    J. CHARDONNE, les Destinées sentimentales, p.
    103.</ex>
    <sem-verbe>perception / réflexe</sem-verbe>
    <pattern-syn>*V* SN</pattern-syn>
    <suppl-syn-sem>(SN_son)</suppl-syn-sem>
  </section>
  ...
</entendre >
```

## Schéma XSD



## Document XML



```
<entendre xmlns:xsi="http://www.w3.org/2001/XMLSchema-
instance"
xsi:noNamespaceSchemaLocation="D:\Documents\xml\ente
ndre_stat.xsd">
...
<section id="entendre_GrandRobert_cit20_291">
  <ex>Dans la lingerie, les femmes repassaient;
[on entendait le bruit des fers] (...) des coups sourds
sur les molletons (...)
J. CHARDONNE, les Destinées sentimentales, p.
103.</ex>
  <sem-verbe>perception / réflexe</sem-verbe>
  <pattern-syn>*V* SN</pattern-syn>
  <suppl-syn-sem>(SN_son)</suppl-syn-sem>
</section>
...
</entendre>
```

## Valeur informationnelle ajoutée (manuellement)

- **BALISAGE** : attribut *id*
  - entendre\_LaMeuse\_021001\_04
  - entendre\_GrandRobert\_cit20\_291

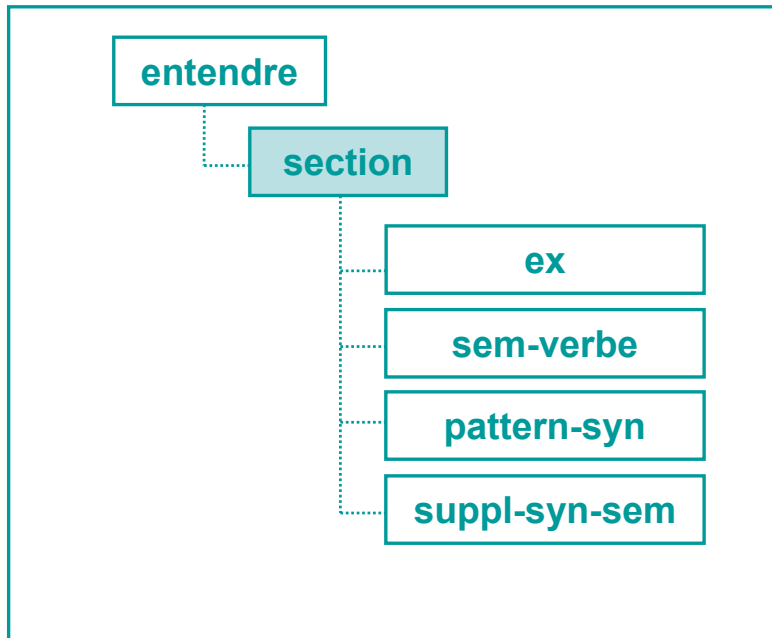


Schéma XSD

```
<section id="entendre_GrandRobert_cit20_291">  
  <ex>Dans la lingerie, les femmes repassaient;  
  [on entendait le bruit des fers] (...) des coups sourds  
  sur les molletons (...)  
  J. CHARDONNE, les Destinées sentimentales, p.  
  103.</ex>  
  <sem-verbe>perception / réflexe</sem-verbe>  
  <pattern-syn>*V* SN</pattern-syn>  
  <suppl-syn-sem>(SN_son)</suppl-syn-sem>  
</section>
```

Document XML

## Valeur informationnelle ajoutée (manuellement)

- IDENTIFICATION DE LA CITATION : `<ex>`, [ ]

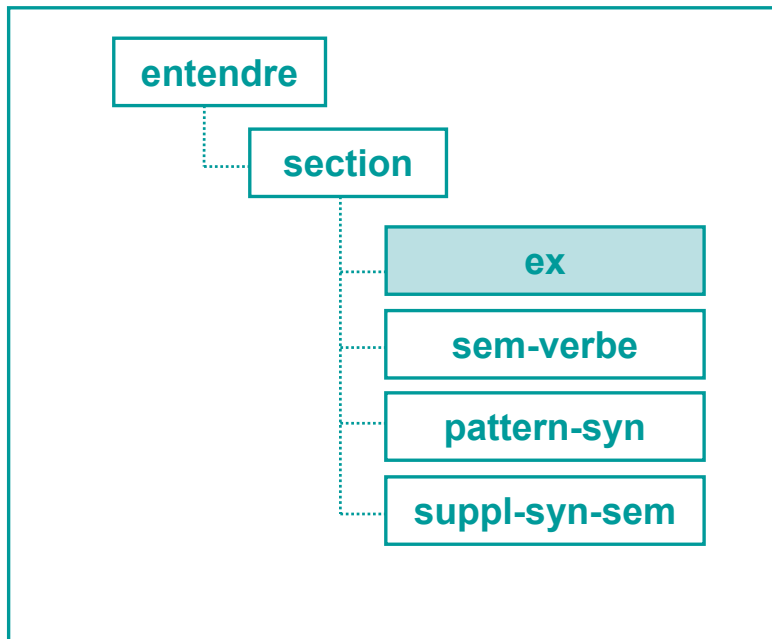


Schéma XSD

```
<section id="entendre_GrandRobert_cit20_291">
  <ex> Dans la lingerie, les femmes repassaient;
  [on entendait le bruit des fers] (...) des coups sourds
  sur les molletons (...)
  J. CHARDONNE, les Destinées sentimentales, p.
  103.</ex>
  <sem-verbe>perception / réflexe</sem-verbe>
  <pattern-syn>*V* SN</pattern-syn>
  <suppl-syn-sem>(SN_son)</suppl-syn-sem>
</section>
```

Document XML

## Valeur informationnelle ajoutée (manuellement)

- SENS DU VERBE DANS LA CITATION : <sem-verbe>

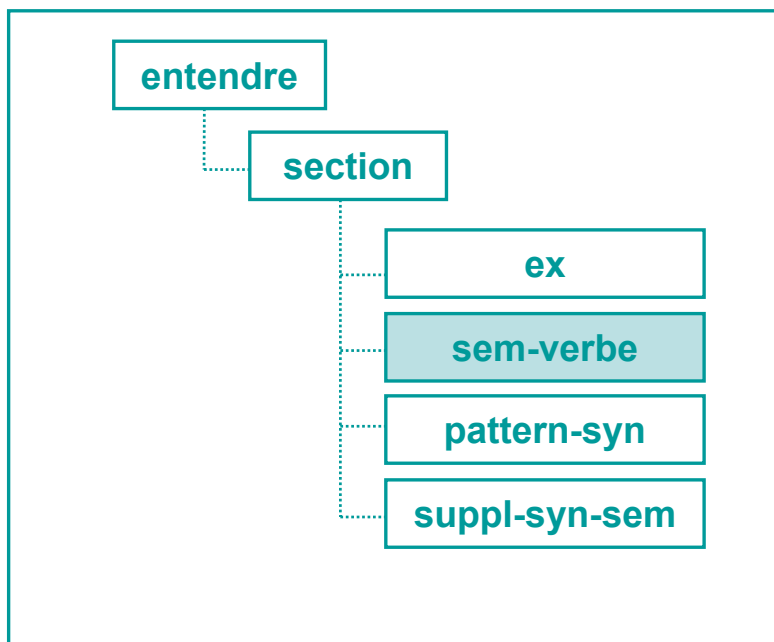


Schéma XSD

```
<section id="entendre_GrandRobert_cit20_291">  
  <ex>Dans la lingerie, les femmes repassaient;  
  [on entendait le bruit des fers] (...) des coups  
  sourds sur les molletons (...)  
  J. CHARDONNE, les Destinées sentimentales, p.  
  103.</ex>  
  <sem-verbe>perception / réflexe</sem-verbe>  
  <pattern-syn>*V* SN</pattern-syn>  
  <suppl-syn-sem>(SN_son)</suppl-syn-sem>  
</section>
```

Document XML

# La polysémie

MOT x

MOT y

entendre

MOT z

MOT w

liste des sens issue

- des notices des dictionnaires (étude des champs sémantiques)
- du travail sur le corpus

Compétence auditive  
Perception auditive  
Attention auditive  
Écoute officielle  
Écoute favorable  
Mémoire auditive  
Transgression modale  
Accepter  
Juger  
Perception d'une situation  
Vouloir  
Relation - accord  
Relation réciproque  
Compréhension  
Habilité

## Valeur informationnelle ajoutée (manuellement)

- SENS DU VERBE DANS LA CITATION : <sem-verbe>

Exemple 1. entendre\_LaMeuse\_221001\_12

L'Abbaye avait été reprise dès 1999 par M. Etienne Jonniaux  
[ qui **entendait** l'exploiter. ] = vouloir

Exemple 2. entendre\_LaMeuse\_251001\_21

[ Les conseillers **ont** ensuite **entendu**  
le président du CPAS, Willy Willock, présenter sa modification  
budgétaire. ] = perception / attention



## Valeur informationnelle ajoutée (manuellement)

- ANNOTATION GRAMMATICALE PARTIELLE : <pattern-syn>

POS, PS – EAGLES 1996

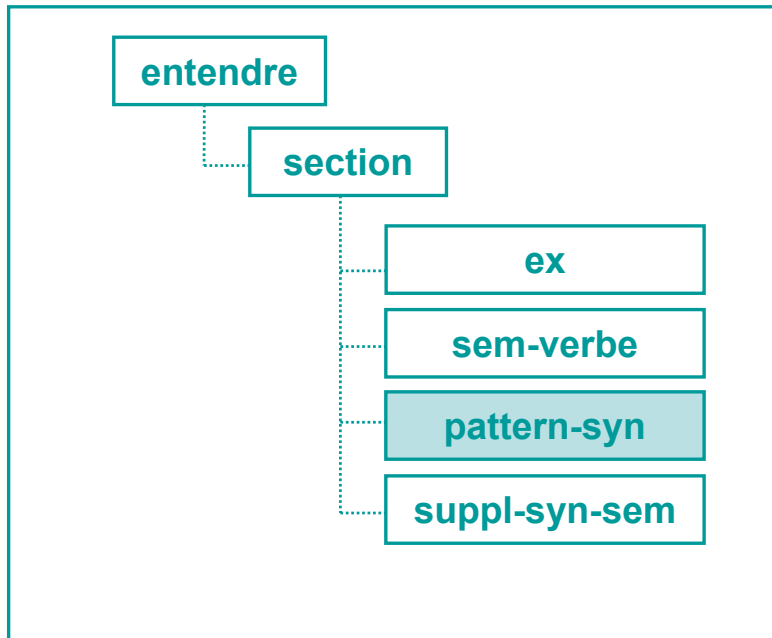


Schéma XSD

```
<section id="entendre_GrandRobert_cit20_291">
  <ex>Dans la lingerie, les femmes repassaient;
  [on entendait le bruit des fers] (...) des coups
  sourds sur les molletons (...)
  J. CHARDONNE, les Destinées sentimentales, p.
  103.</ex>
  <sem-verbe>perception / réflexe</sem-verbe>
  <pattern-syn>*V* SN</pattern-syn>
  <suppl-syn-sem>(SN_son)</suppl-syn-sem>
</section>
```

Document XML

## La construction syntaxique

MOT x

MOT y

entendre

MOT z

MOT w

Liste des patrons  
syntaxiques

\*V\* SN  
[\*V\* [V]]  
[\*V\* [x V]]  
[\*V\* Compl]  
[\*V\* P]  
[AuxPass \*V\*]  
[CLse \*V\*]  
[\*V\*]  
[CLind \*V\*]

## Valeur informationnelle ajoutée (manuellement)

- ANNOTATION GRAMMATICALE PARTIELLE <pattern-syn>

Exemple 1. entendre\_LaMeuse\_221001\_12

L'Abbaye avait été reprise dès 1999 par M. Etienne Jonniaux

[ qui **entendait** l' **exploiter.** ]  
= [ \*V\* [CLdir V ] ]

Exemple 2. entendre\_LaMeuse\_251001\_21

[ Les conseillers **ont** ensuite **entendu**  
**le président du CPAS, Willy Willock, présenter sa modification**  
**budgétaire** ]  
= [ \*V\* [SN V ...] ]

## Valeur informationnelle ajoutée (manuellement)

- ANNOTATIONS SUPPLÉMENTAIRES : `<suppl-syn-sem>`  
catégorie sémantique ou syntaxique de certains constituants

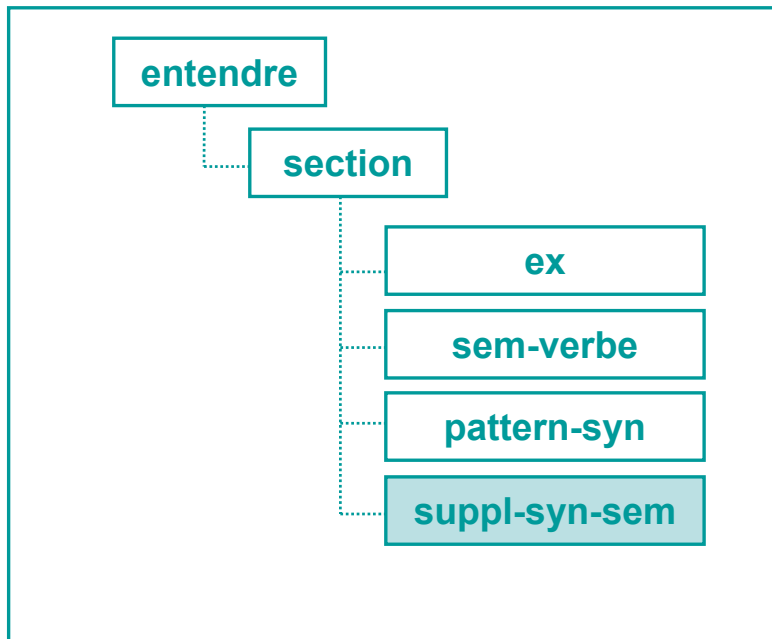


Schéma XSD

```
<section id="entendre_GrandRobert_cit20_291">
  <ex>Dans la lingerie, les femmes repassaient;
  [on entendait le bruit des fers] (...) des coups
  sourds sur les molletons (...)
  J. CHARDONNE, les Destinées sentimentales, p.
  103.</ex>
  <sem-verbe>perception / réflexe</sem-verbe>
  <pattern-syn>*V* SN</pattern-syn>
  <suppl-syn-sem>(SN_son)</suppl-syn-sem>
</section>
```

Document XML

## Valeur informationnelle ajoutée (manuellement)

- ANNOTATIONS SYNTAXIQUES SUPPLÉMENTAIRES : <suppl-syn-sem>

Exemple 3. entendre\_LaMeuse\_161101\_14

ASBL présentera son activité sur le site internet.

[ Dehors, du bruit **se fait entendre.** ] Quelques "monstres" affamés vont bientôt rejoindre leurs parents à

<pattern-syn> [CLse \*V\*] </pattern-syn>

<suppl-syn-sem> AuxMod = faire </suppl-syn-sem>

## Valeur informationnelle ajoutée (manuellement)

- ANNOTATIONS SÉMANTIQUES SUPPLÉMENTAIRES : <suppl-syn-sem>

Exemple 4. entendre\_GrandRobert\_cit20\_472

La place était très éclairée. Jenny, hardiment, se jeta au milieu des voitures. Lui aussi. Il évita de justesse un autobus, et [ entendit les injures du chauffeur.]

Martin du Gard, *les Thibault*, t. VI, p. 49.

<pattern-syn> [\*V\* SN] </pattern-syn>

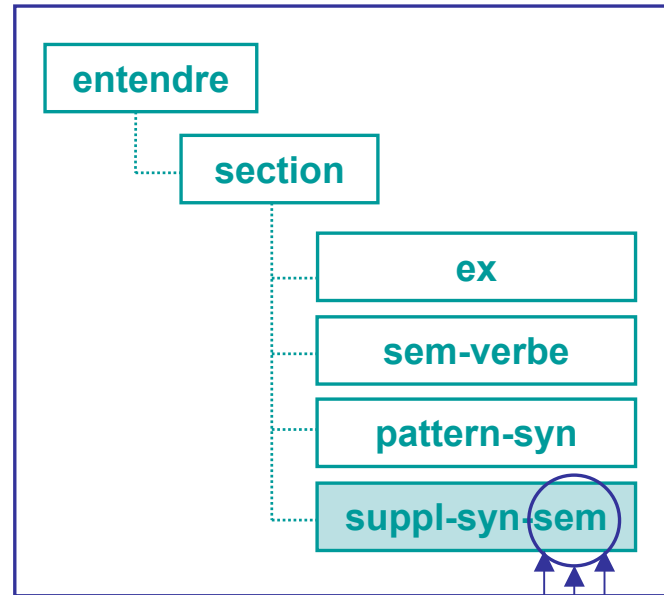
<suppl-syn-sem> SN\_parole </suppl-syn-sem>

Exemple 5. entendre\_LaMeuse\_251001\_21

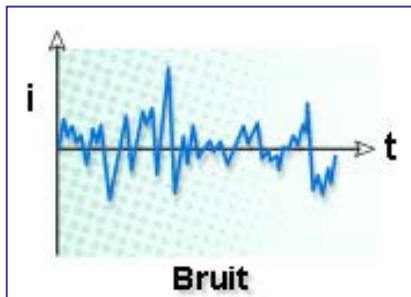
[ Les conseillers ont ensuite entendu le président du CPAS, Willy Willock, présenter sa modification budgétaire. ]

<pattern-syn> [\*V\* [SN V ...]] </pattern-syn>

<suppl-syn-sem> sujet\_humain </suppl-syn-sem>

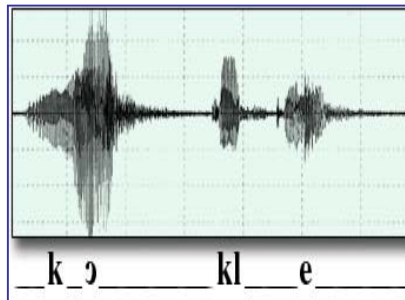


## • MODE ENVIRONNEMENTAL



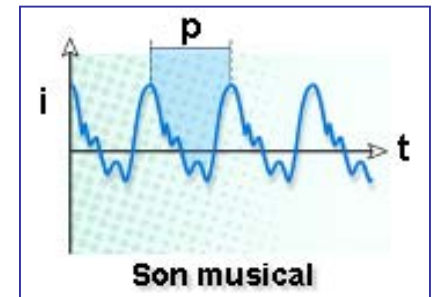
<http://www.iurc.montp.inserm.fr/cric/audition/fran%E7ais/sound/fsound.htm>

## • MODE PHONÉTIQUE

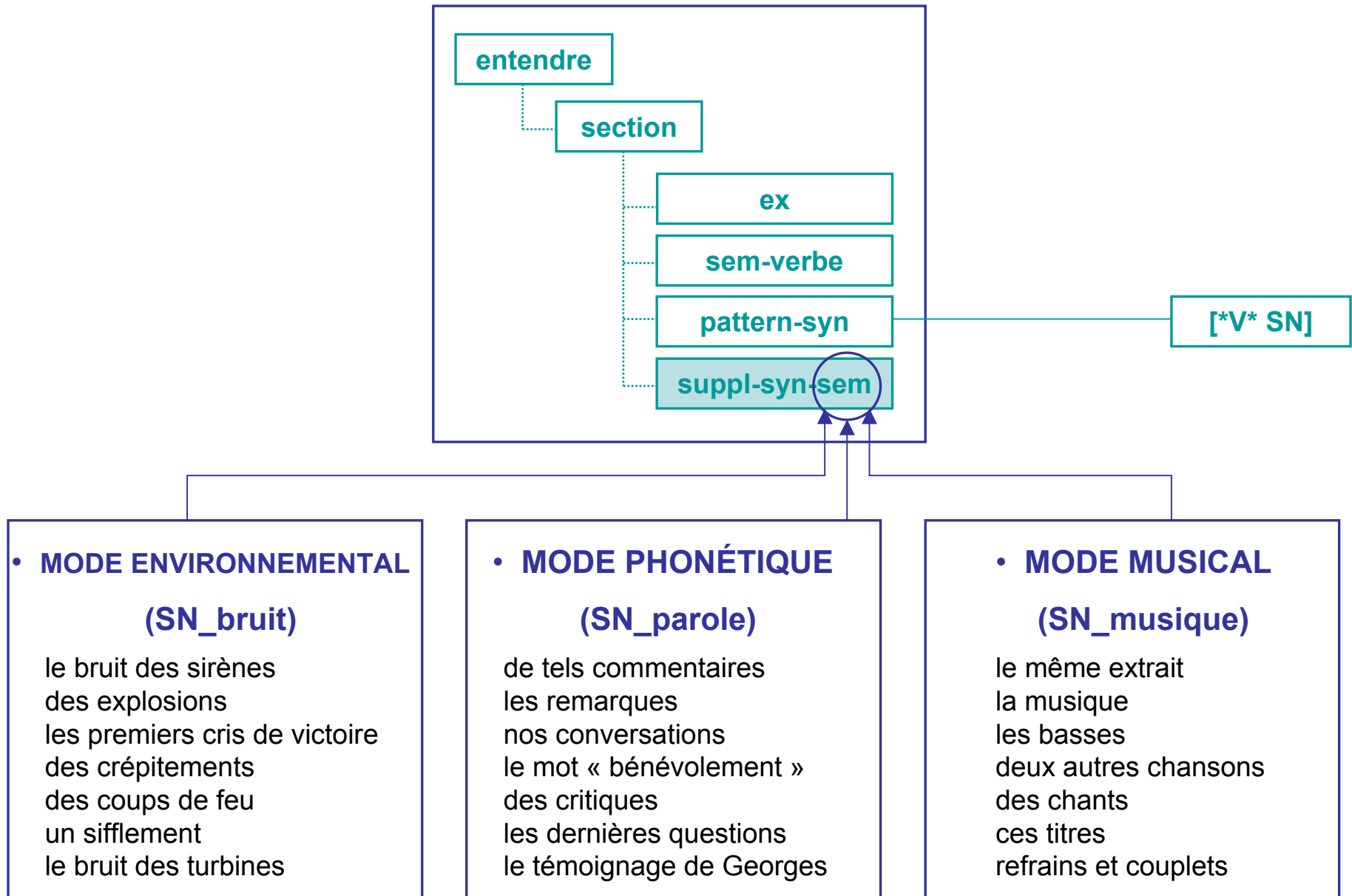


<http://www.iurc.montp.inserm.fr/cric/audition/fran%E7ais/sound/fsound.htm>

## • MODE MUSICAL



<http://www.iurc.montp.inserm.fr/cric/audition/fran%E7ais/sound/fsound.htm>





## Sections du corpus

```
<section id="entendre_GrandRobert_cit20_001">  
  <ex>...</ex>  
  <sem-verbe>...</sem-verbe>  
  <pattern-syn>...</pattern-syn>  
  <suppl-syn-sem>...</suppl-syn-sem>  
</section>
```

...

```
<section id="entendre_GrandRobert_cit20_720">...  
</section>
```

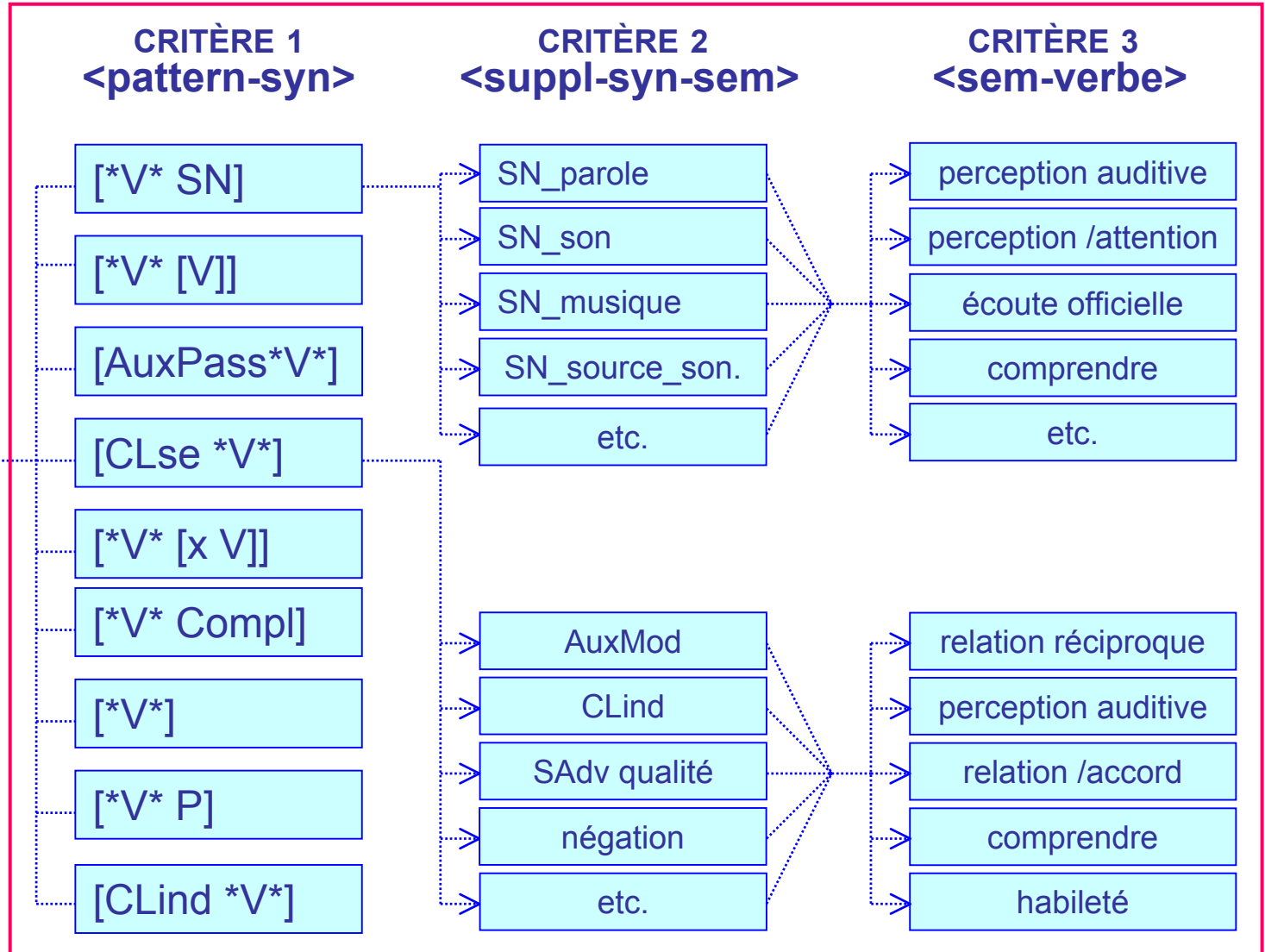
\*V\*:entendre

```
<section id="entendre_LaMeuse_021001_01">  
  <ex>...</ex>  
  <sem-verbe>...</sem-verbe>  
  <pattern-syn>...</pattern-syn>  
  <suppl-syn-sem>...</suppl-syn-sem>  
</section>
```

...

```
</section id="entendre_LaMeuse_070102_10">...  
</section>
```

## Classement des sections de chaque sous-corpus



# Développement de la recherche

## Domaine : sciences cognitives

AXE COGNITIF

AXE LINGUISTIQUE

PSYCHOLOGIE EXPÉR.

LINGUISTIQUE

•fondements neuro-physiologiques

•fondements neuronaux

•cognition sensorielle

fondements historiques

champs sémantiques

étude de corpus

## Domaine : informatique

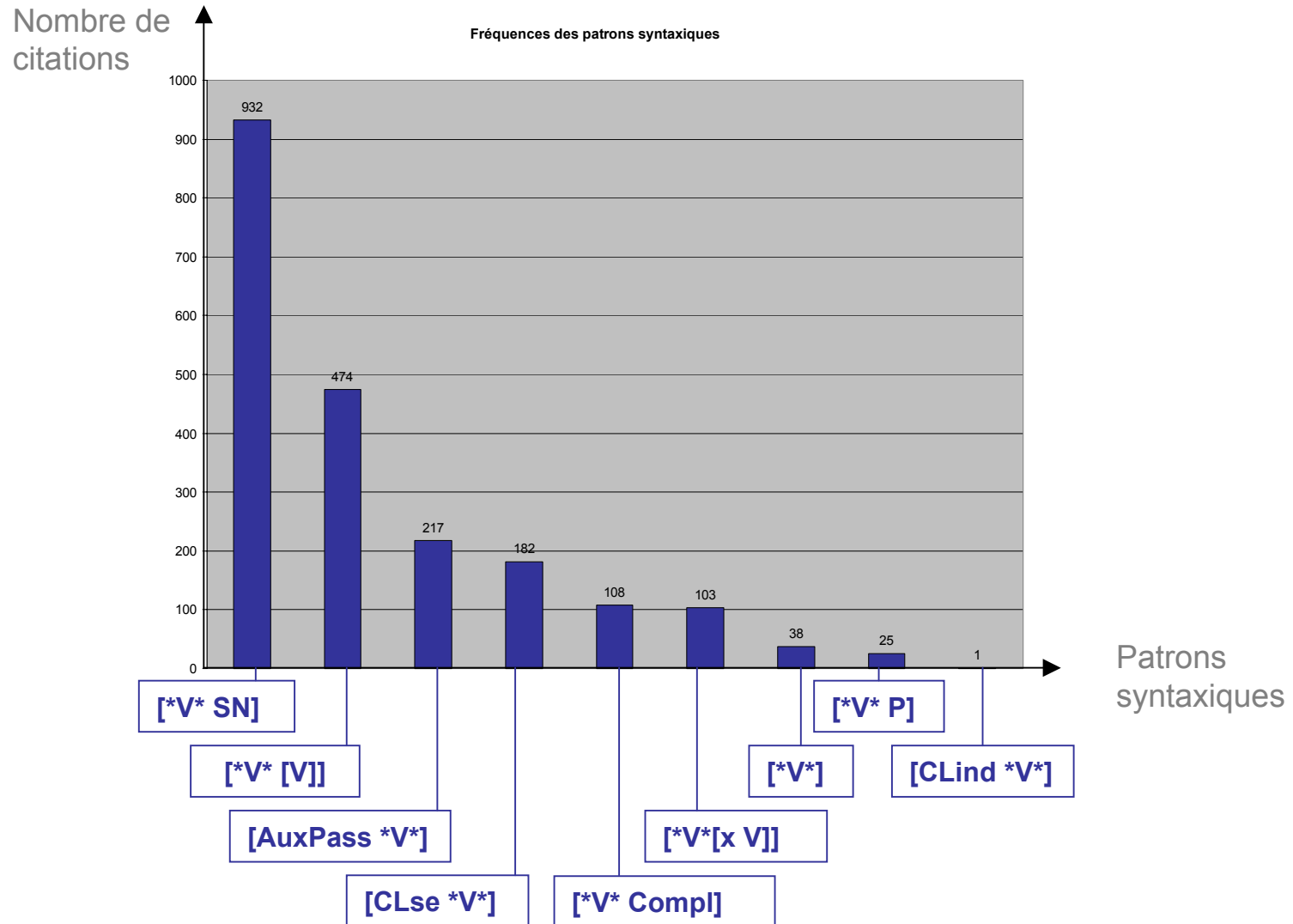
AXE INFORMATIQUE

FORMALISATION

APPLICATION

corpus XML

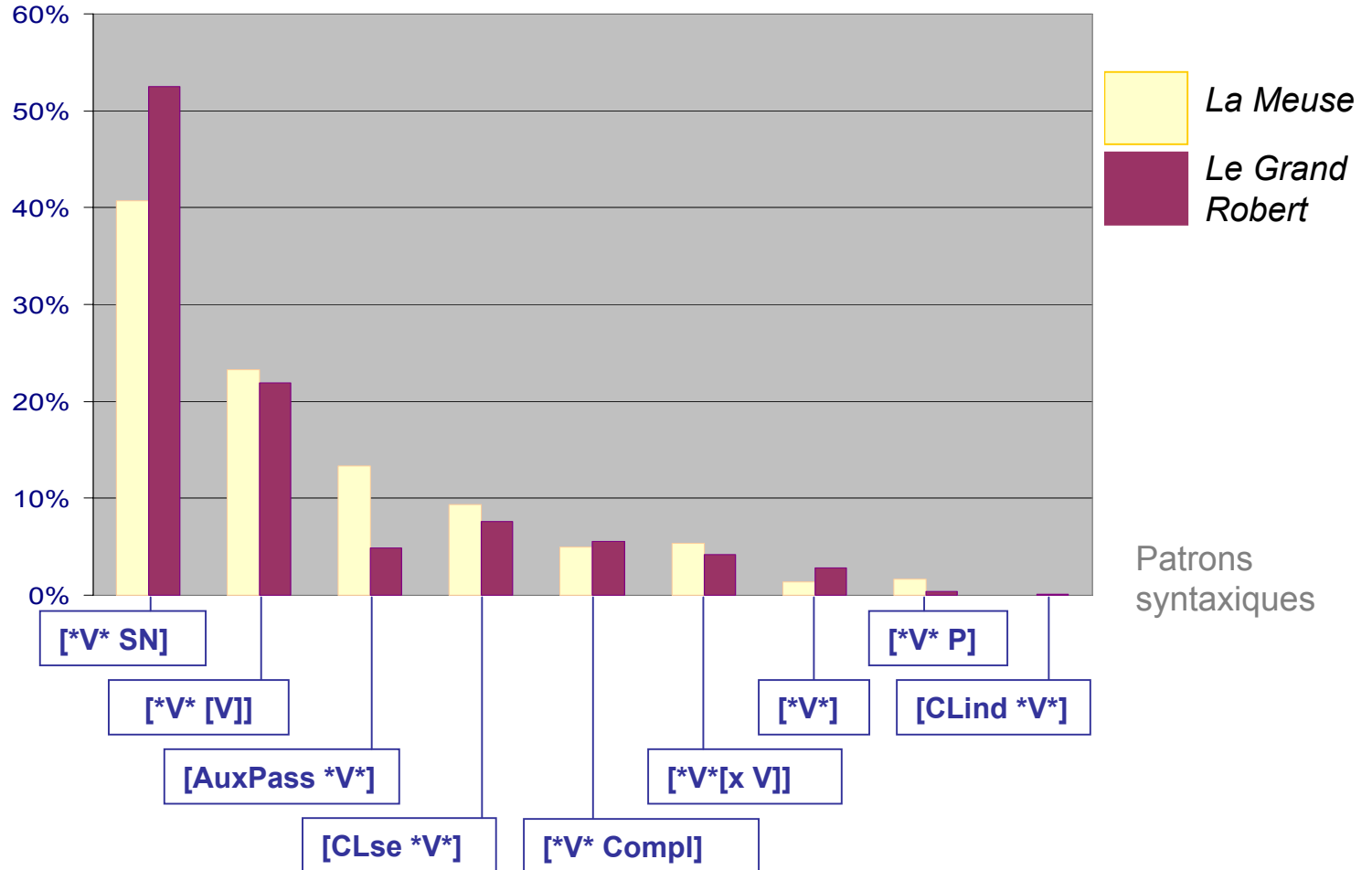
## Fréquence des patrons syntaxiques (verbe *entendre*)



# Axe linguistique – étude de corpus

Les fréquences, en pourcentage, des patrons syntaxiques

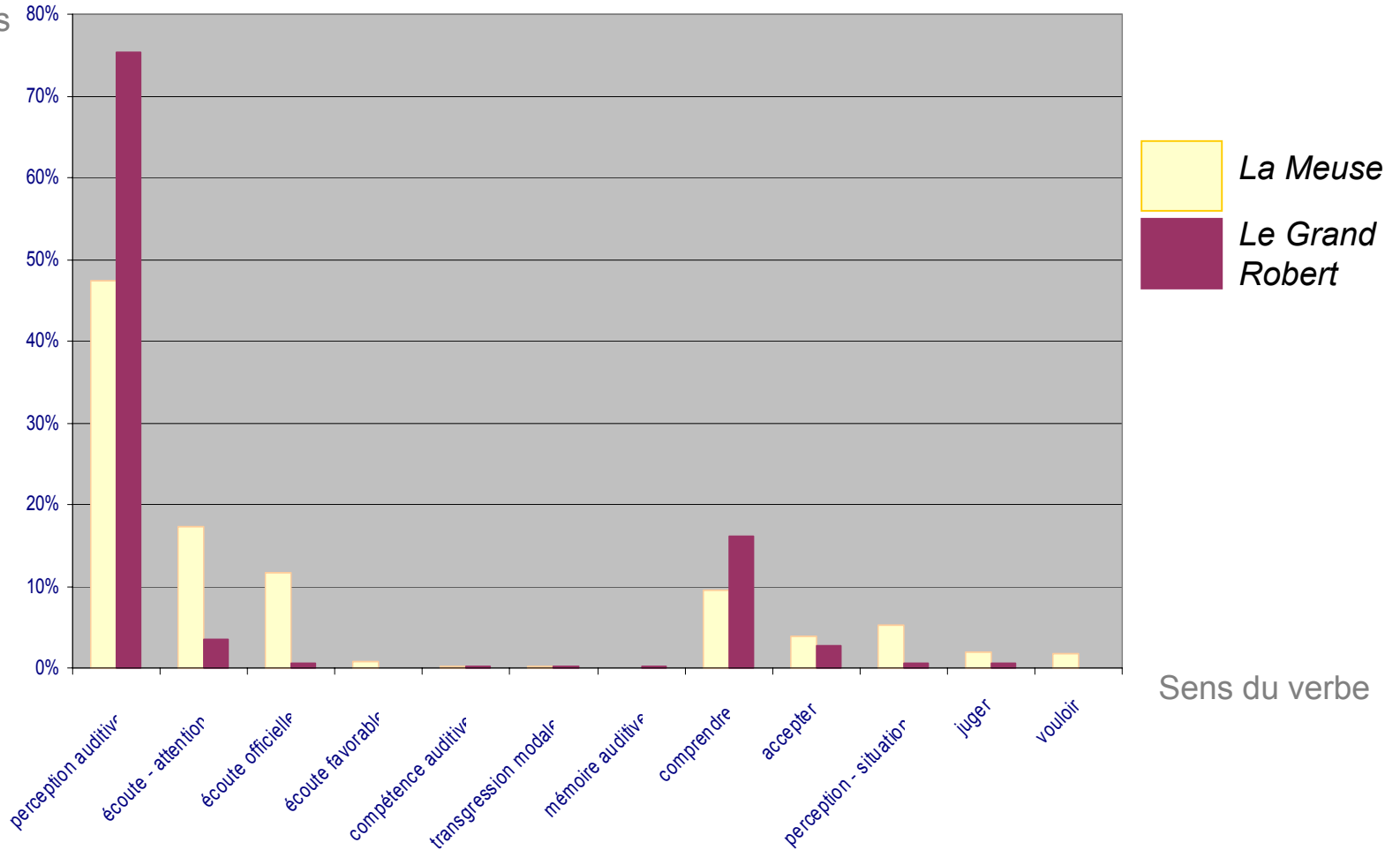
Pourcentage au sein du corpus



# Axe linguistique – étude de corpus

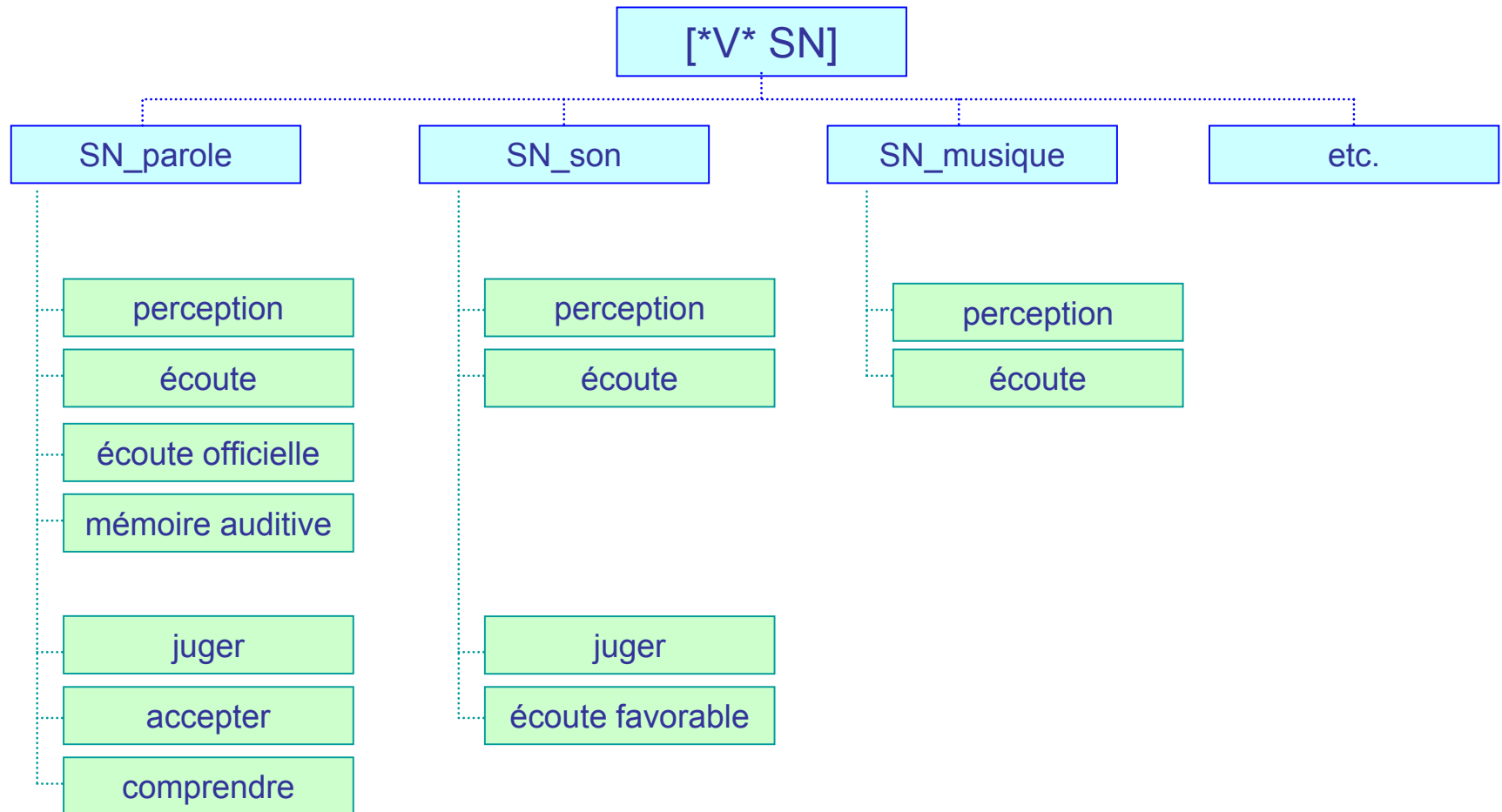
## Les sens du verbe "entendre" en [\*V\* SN]

Pourcentage au sein du corpus

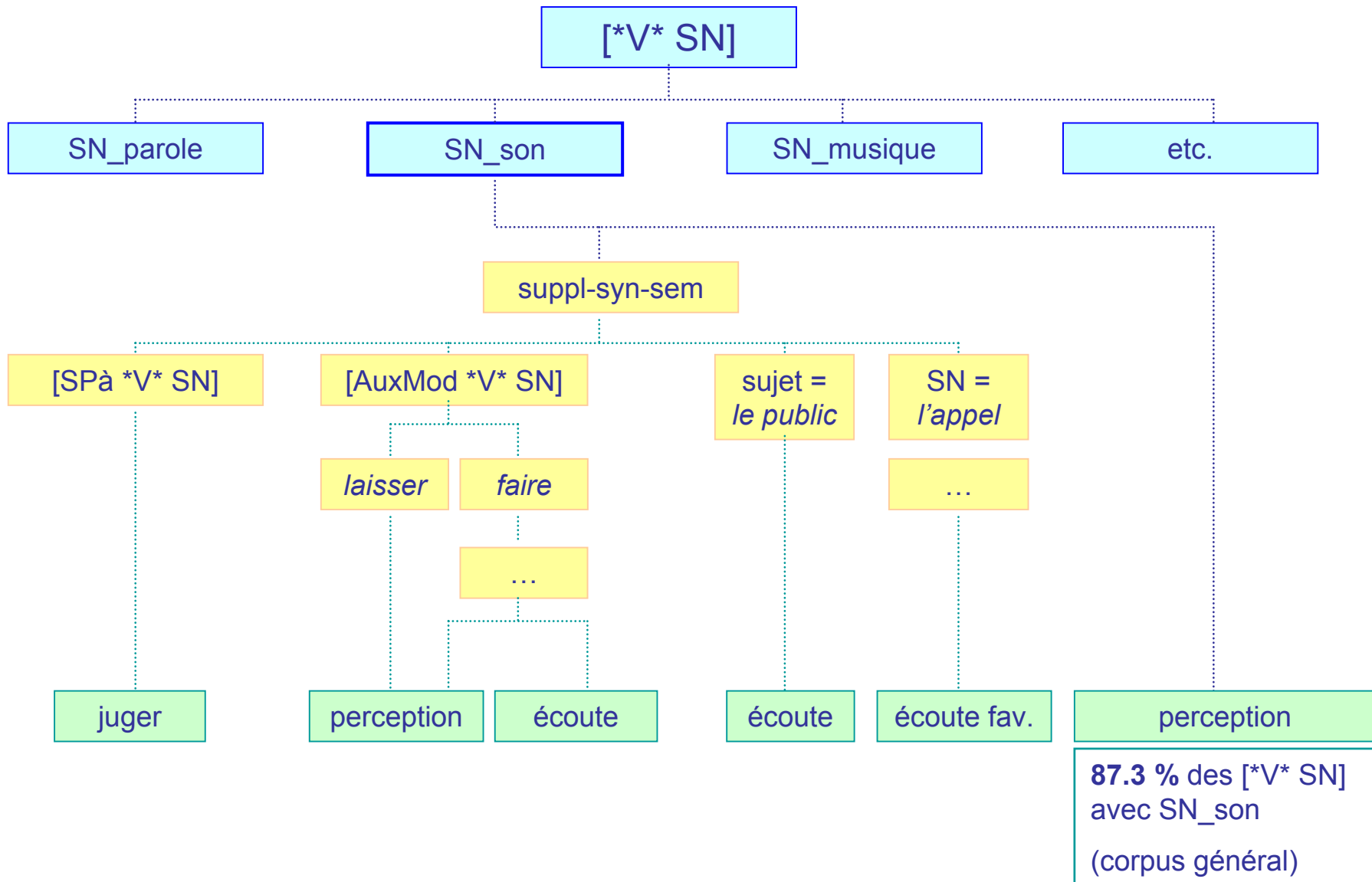


Sens du verbe

# Axe linguistique – étude de corpus – préférences sélectionnelles



# Axe linguistique – étude de corpus – analyse locale





# Développement de la recherche

## Domaine : sciences cognitives

AXE COGNITIF

AXE LINGUISTIQUE

PSYCHOLOGIE EXPÉR.

LINGUISTIQUE

- fondements neuro-physiologiques
- fondements neuronaux
- cognition sensorielle

fondements historiques

champs sémantiques

étude de corpus

analyse linguistique

## Domaine : informatique

AXE INFORMATIQUE

FORMALISATION

APPLICATION

corpus XML

# La polysémie

MOT x

MOT y

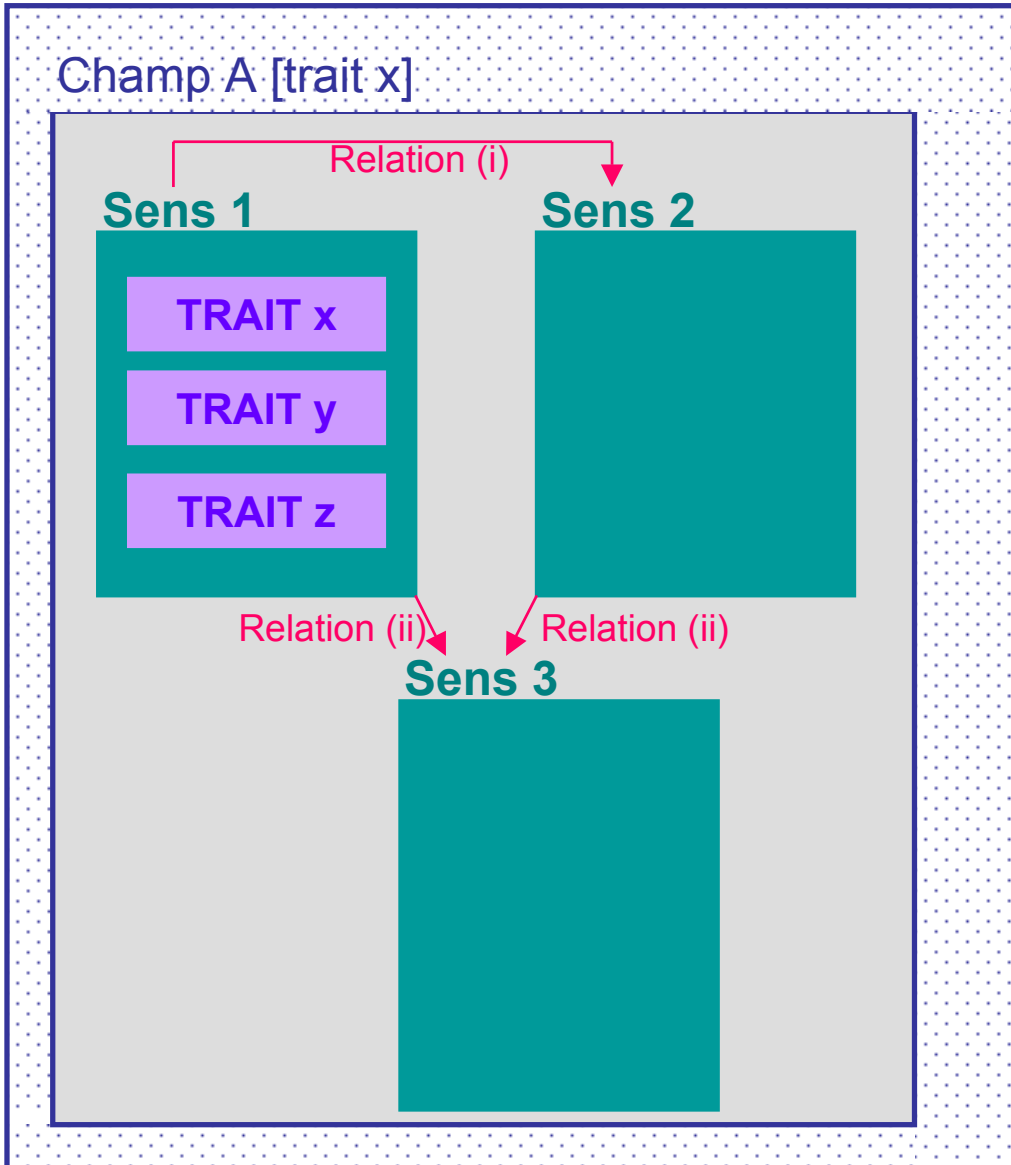
entendre

MOT z

MOT w

Compétence auditive  
Perception auditive  
Attention auditive  
Écoute officielle  
Écoute favorable  
Mémoire auditive  
Transgression modale  
Accepter  
Juger  
Perception d'une situation  
Vouloir  
Relation - accord  
Relation réciproque  
Compréhension  
Habilité

## Verbe



- **Décomposition** des sens en traits sémantiques
- Introduction de la **sémantique cognitive** pour définir ces traits
- Établissement des **relations** entre les différents sens
- Établissement des **champs** et des relations qu'ils entretiennent
- **Comparaison** de la sémantique des différents verbes de perception sensorielle

# Développement de la recherche

## Domaine : sciences cognitives

AXE COGNITIF

AXE LINGUISTIQUE

PSYCHOLOGIE EXPÉR.

LINGUISTIQUE

- fondements neuro-physiologiques
- fondements neuronaux
- cognition sensorielle

fondements historiques

champs sémantiques

étude de corpus

analyse linguistique

## Domaine : informatique

AXE INFORMATIQUE

FORMALISATION

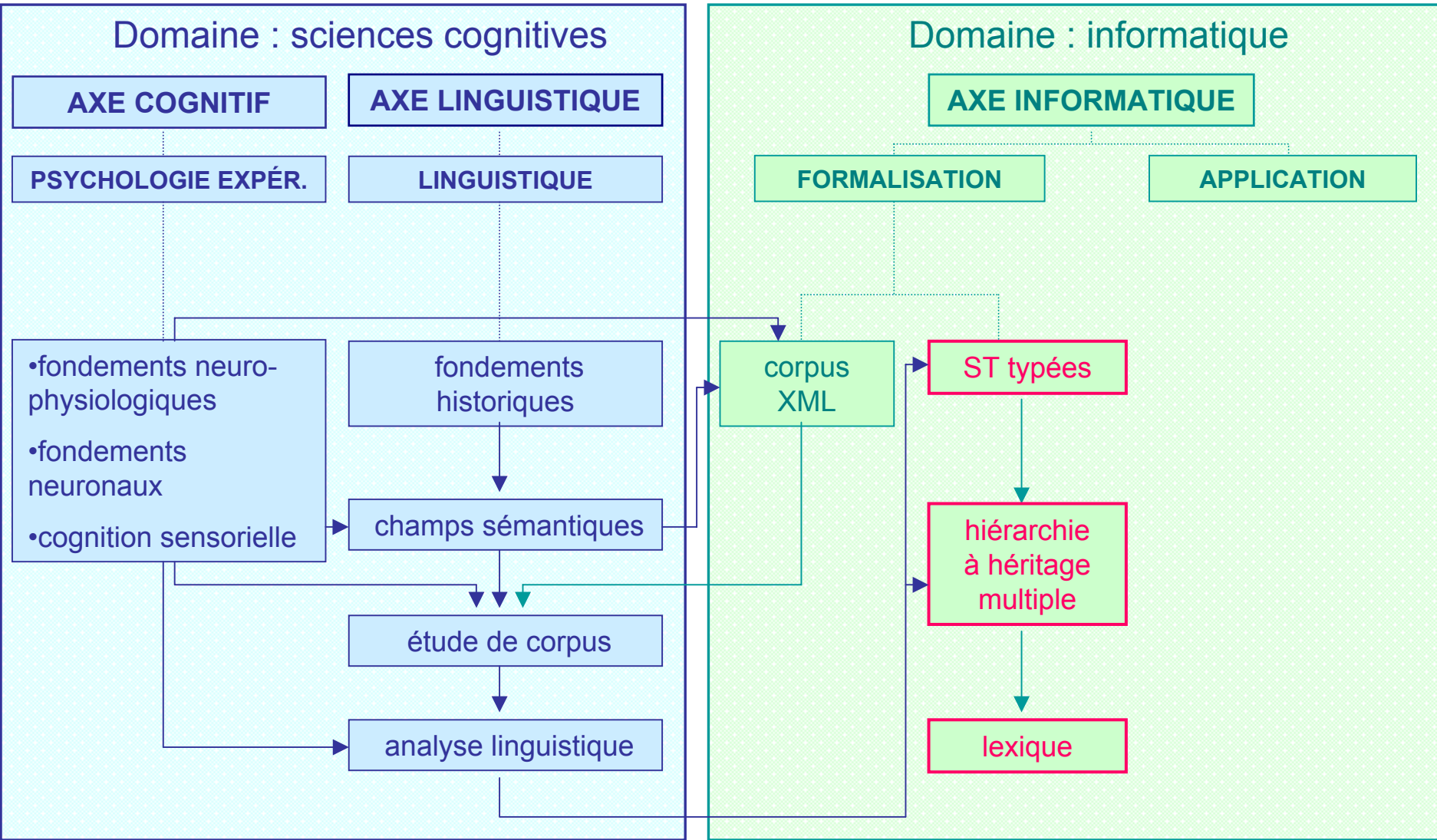
APPLICATION

corpus XML

ST typées

hiérarchie à héritage multiple

lexique



## Problématique

- Coupure entre lexique computationnel et analyse linguistique
- Redondance et énumération dans les lexiques computationnels

<b>Rive<sub>1</sub></b>			<b>Rive<sub>2</sub></b>
CAT	SN	CAT	SN
DEF	bordure_mer/lac	DEF	bord_four

## Objectifs

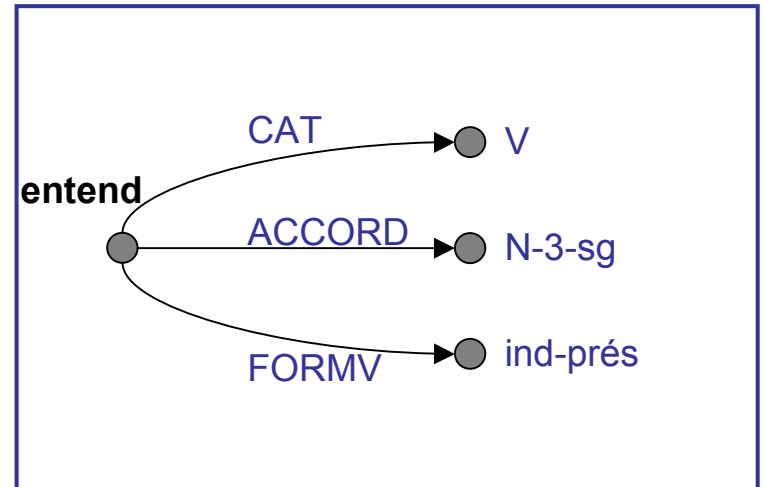
- Transposition de l'analyse linguistique dans le lexique
- Réduire la redondance en élaborant un système de types pour les verbes de perception sensorielle en français

## Structures de traits (ST) typées

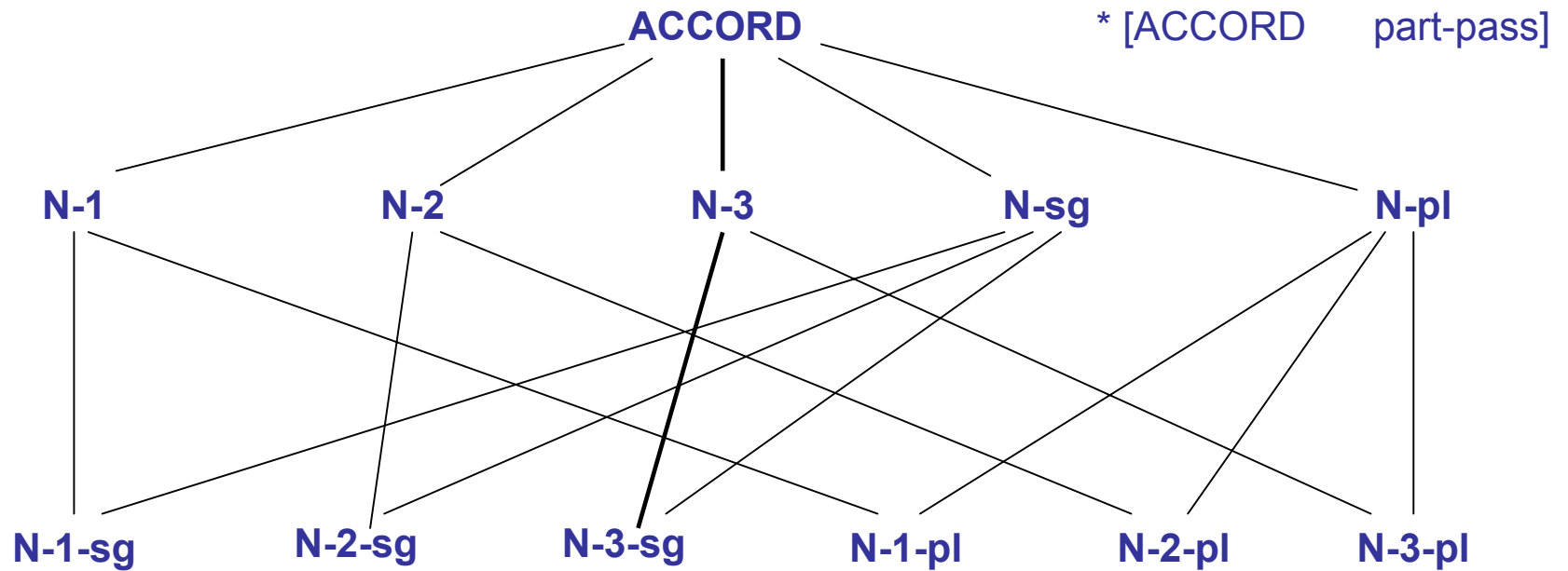
Matrice d'attributs-valeurs

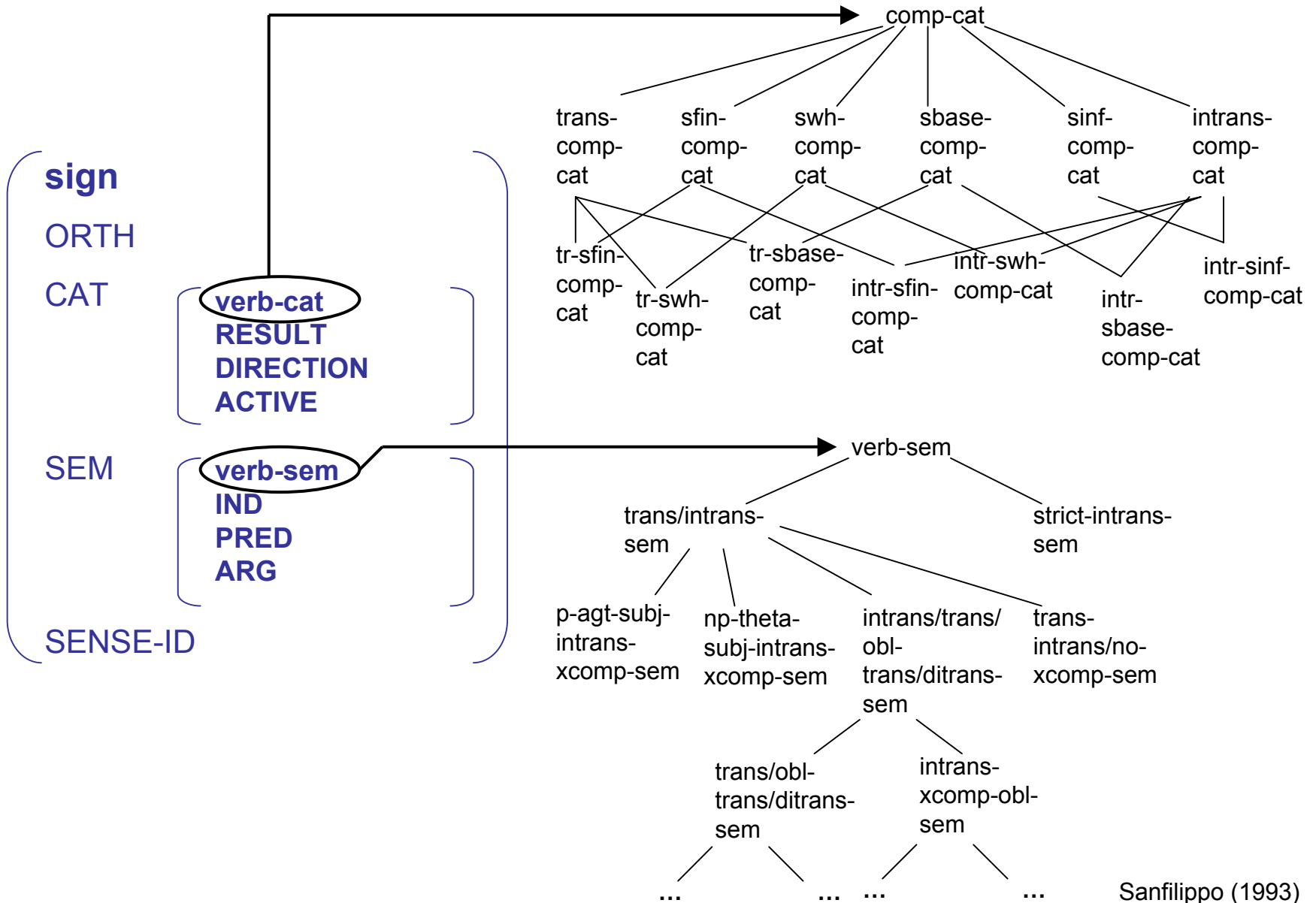
<b>entend</b>	
CAT	V
ACCORD	N-3-sg
FORMV	ind-prés

Graphe acyclique orienté



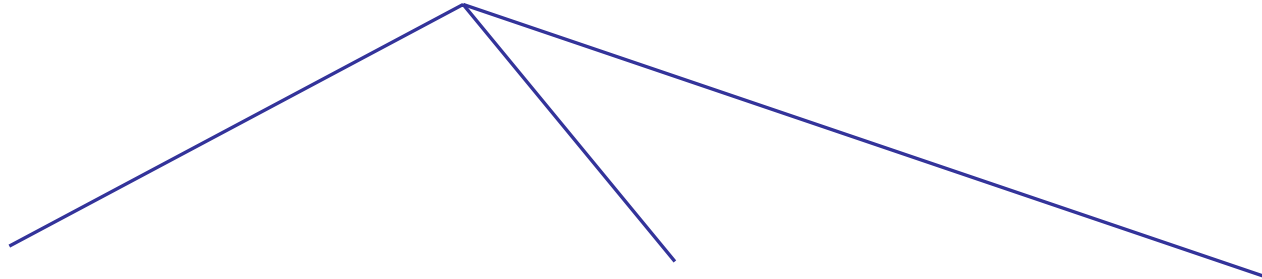
## Organisation des types en treillis







Méta-entrée = un type lexical



**Sous-type 1** = extension du type  
(un emploi du mot)

**Sous-type 2** = extension du type  
(un emploi du mot)

...

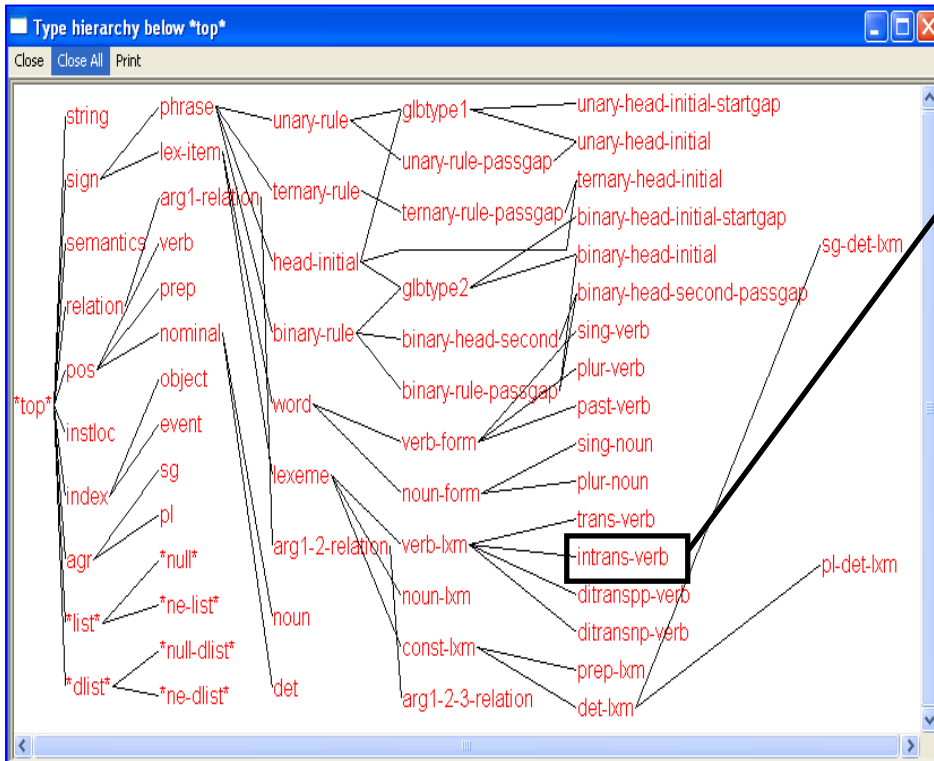
Énumérations « traditionnelles »



## Linguistic Knowledge Building (LKB)

- Environnement de développement de grammaires et de lexiques
- Environnement pour les formalismes à base de contraintes

## Hiérarchie des types



## Définition d'un type

```
intrans-verb - definition
Parents = verb-lxm
[intrans-verb
SPR: [*top*
  FIRST: [*top*
    SEM: [*top*
      INDEX: <0> = *top*]]
  REST: *null*]
COMPS: *null*
SEM: [*top*
  RELS: [*top*
    LIST: [*top*
      FIRST: [arg1-relation
        ARG1: <0>]
      REST: <1> = *top*]
    LAST: <1>]]]]]
```

## Définition d'un type

## Entrée lexicale

```
intrans-verb - definition
Close Close All Print

intrans-verb - definition

Parents = verb-lxm

intrans-verb
SPR: [*top*
  FIRST: [*top*
    SEM: [*top*
      INDEX: <0> = *top*]]
  REST: *null*]
COMPS: *null*
SEM: [*top*
  RELS: [*top*
    LIST: [*top*
      FIRST: [arg1-relation
        ARG1: <0>]
      REST: <1> = *top*]
    LAST: <1>]]]]
```

```
bark - expanded
Close Close All Print

bark - expanded

intrans-verb
ORTH: [*dlist*
  LIST: [*ne-list*
    FIRST: bark
    REST: <0> = *list*]
  LAST: <0>]
HEAD: [verb
  MOD: *null*]
SPR: [*ne-list*
  FIRST: phrase
  REST: *null*]
COMPS: *null*
SEM: [semantics
  INDEX: <1> = [event
    INSTLOC: instloc]
  RELS: [*dlist*]
GAP: [*dlist*
  LIST: <2> = *list*
  LAST: <2>]
ARGS: *list*]
```

# Développement de la recherche - conclusion

## Domaine : sciences cognitives

AXE COGNITIF

AXE LINGUISTIQUE

PSYCHOLOGIE EXPÉR.

LINGUISTIQUE

- fondements neuro-physiologiques
- fondements neuronaux
- cognition sensorielle

fondements historiques

champs sémantiques

étude de corpus

analyse linguistique

## Domaine : informatique

AXE INFORMATIQUE

FORMALISATION

APPLICATION

corpus XML

ST typées

lexique dans LKB

hiérarchie à héritage multiple

lexique

