Collecting terms and variants of terms fitting in a multilingual framework

Béatrice DAILLE

LS2N, University of Nantes, France

July 21st, 2017
Outline

1 Introduction
Outline

1. Introduction

2. Characterization
Outline

1 Introduction

2 Characterization

3 Linguistic processes leading to variants
   - Denominative variants
   - Conceptual variants

4 Variant discovery
   - Syntagmatic analysis
   - Distributional analysis
   - Inference rules

5 Software

6 Conclusion and perspectives
Outline

1. Introduction

2. Characterization

3. Linguistic processes leading to variants
   - Denominative variants
   - Conceptual variants

4. Variant discovery
   - Syntagmatic analysis
   - Distributional analysis
   - Inference rules
Outline

1 Introduction

2 Characterization

3 Linguistic processes leading to variants
   - Denominative variants
   - Conceptual variants

4 Variant discovery
   - Syntagmatic analysis
   - Distributional analysis
   - Inference rules

5 Software
Outline

1 Introduction

2 Characterization

3 Linguistic processes leading to variants
   - Denominative variants
   - Conceptual variants

4 Variant discovery
   - Syntagmatic analysis
   - Distributional analysis
   - Inference rules

5 Software

6 Conclusion and perspectives
Introduction

Automatic detection of term variant
## Automatic detection of term variant

- Symbolic approaches
- Empirical approaches
Introduction

Automatic detection of term variant
- Symbolic approaches
- Empirical approaches

Issues
- Variants have many forms
- Each form has very few occurrences

Many methods are required.
Assumption

Languages

French, English, German, Spanish and Russian.

these languages are subject to the same variation processes of variation and can be formally described.
The definition of variant

**Definition**

[Daille et al., 1996:201]
A variant of a term is an utterance which is semantically and conceptually related to an original term.

**Terms**

- simple or complex terms
- morphological compounds: native or neoclassical
- syntagmatic compounds
Organisation of variants

- **Denominative variants**: to respond to the properties of transparency and of minimality of the denominative core of the term;
- **Conceptual variants**: to anchor the term in the system of knowledge instantiated in the text;
- **Linguistic variants**: to link the term into the language system only.

Can be paired with Freixa’s categories of variants [Freixa, 2006]
Denominative variants

Definition

Denominative variants reflect a synonymy relation.

exact synonyms or approximate synonyms
lexicalised forms
(MED) En: histamine flare test → histamine test [Collet, 1997]
paraphrases
(AGR) Fr: protéine végétale ’plant protein’ → protéine d’origine végétale ’protein from plant’
Conceptual variants

**Definition**

Conceptual variants reflect a conceptual or a semantic relation.

**antonymy, taxonomy, meronymy and complex relations**

**lexicalisable forms**

(MED) En: *blood cell* → *blood cell line*

**lexicalised forms**

(MED) En: *blood cell* → *cell* [Kister, 2000]
Relationships between categories of variants

- Conceptual variants
  - Denominative variants
    - Linguistic variants
    - Linguistic variants
  - Linguistic variants
- Denominative variants
  - Linguistic variants
- Base-term
Property of variants

1. a variant always involves at least one term;
2. a variant is obtained by applying at least one linguistic operation;
3. a term can produce several variants;
4. the number or utterances of the term in a text is slightly superior to the number of utterances of the variant.
Denominative variants

1 synonymic substitution (lexical content)
2 competing patterns
3 simplification (minimality criterion)
4 exemplification (transparency criterion)
Synonymic substitution (1)

Morpheme

several affixes in competition
(decision-making process) Fr: -eur/-aire for nouns
décideur ↔ décisionnaire 'decision maker'
(decision-making process) Fr: -el/-aire for adjectives
processus décisionnel ↔ processus décisionnaire 'decision-making process'

Functional word

several prepositions in competition for N P N
(AGR) Fr: chromatographie en colonne ↔ chromatographie sur colonne
'column chromatography'
Synonymic lexical substitution (2)

Morphological compounds
several roots in competition
(Nervous system) Fr: neurologie
(Epidermis and dermis) Fr: névrodermite
several lexemes in competition
(EOL) De: Rotorblattprofil ’blade profile’ ↔ Flügelprofile ’wing profile’

Syntagmatic compounds
several lexemes in competition
(MED) Fr: parenthèse thérapeutique ’therapeutic range’ ↔ fenêtre thérapeutique ’therapeutic window’
Conceptual variants

1 Expansion
2 Anaphorical reduction
## Expansion (1)

### Derivation

(EOL) **Fr:** éolien/A ’eolian’ $\rightarrow$ **proéolien/A** ’pro-eolian’

### Predication

(DIA) **En:** sentinel node $\rightarrow$ sentinel node biopsy
Expansion: Modification (2)

**Juxtaposition: initial position**

\[ N_2 + N_1 \rightarrow N + N_2 + N_1 \]

(EOL) De: *Windpark* 'wind farm' \( \rightarrow \) *Meerwindpark* 'marine wind farm'

\[ N_2 \ N_1 \rightarrow A \ N_2 \ N_1 \]

(SAT) En: *telecommunication satellite* \( \rightarrow \) *geostationary telecommunication satellite*

**Juxtaposition: post-position**

\[ N_1 \ A_2 \rightarrow N_1 \ A_2 \ A \]

(SAT) Fr: *station terrienne* 'earth station' \( \rightarrow \) *station terrienne brouilleuse* 'interfering earth station'

\[ N_1 \ A_2 \rightarrow N_1 \ A_2 \ P \ A \ N \]

(Transportation) Fr: *transport terrestre* 'ground transportation' \( \rightarrow \) *transport terrestre à grande vitesse* 'high-speed ground transportation'
Expansion: Modification (3)

### Insertion

<table>
<thead>
<tr>
<th>N₁ A₂ → N₁ A/E+A₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>(EOL) Es: acoplamiento dinámico → acoplamiento aerodinámico</td>
</tr>
<tr>
<td>N₁ A₁ → N₁ A A₁</td>
</tr>
<tr>
<td>(EOL) Fr: parc marin 'marine park’ → parc naturel marin 'natural marine park’</td>
</tr>
<tr>
<td>N₁ A₂ P₃ N₄ → N₁ A₂ A P₃ N₄</td>
</tr>
<tr>
<td>(SAT) Fr: service de radiodiffusion par satellite ’fixed-satellite service’ → services communautaires de radiodiffusion par satellite ’domestic fixed-satellite service’</td>
</tr>
</tbody>
</table>
Syntagmatic analysis

**Structural rules**

Term: $X_1 X_2 \rightarrow$ Variant: $X_1 X X_2$

**Linguistic preprocessing**

- part-of-speech tagging
- lemmatisation
- splitting
### Grammar of variants

<table>
<thead>
<tr>
<th>Term</th>
<th>Conceptual variant</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Insertion</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>N/E&lt;sub&gt;1&lt;/sub&gt;+N&lt;sub&gt;2&lt;/sub&gt;</td>
<td>N/E&lt;sub&gt;1&lt;/sub&gt;+N&lt;sub&gt;2&lt;/sub&gt;</td>
</tr>
<tr>
<td>S</td>
<td>A&lt;sub&gt;1&lt;/sub&gt; N&lt;sub&gt;1&lt;/sub&gt;</td>
<td>A&lt;sub&gt;1&lt;/sub&gt; A A N&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
<tr>
<td><strong>Expansion</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>N/N&lt;sub&gt;1&lt;/sub&gt;+N&lt;sub&gt;2&lt;/sub&gt;</td>
<td>N/N+&lt;sub&gt;1&lt;/sub&gt;+N&lt;sub&gt;2&lt;/sub&gt;</td>
</tr>
<tr>
<td>S</td>
<td>N&lt;sub&gt;2&lt;/sub&gt; N&lt;sub&gt;1&lt;/sub&gt;</td>
<td>A N&lt;sub&gt;2&lt;/sub&gt; N&lt;sub&gt;1&lt;/sub&gt;</td>
</tr>
</tbody>
</table>
Features of the grammar of variants

<table>
<thead>
<tr>
<th></th>
<th>De</th>
<th>En</th>
<th>Es</th>
<th>Fr</th>
<th>Ru</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP</td>
<td>M</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>AR</td>
<td>M</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>8</td>
<td>17</td>
<td>17</td>
<td>13</td>
</tr>
<tr>
<td>CT</td>
<td>M</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>11</td>
<td>9</td>
<td>15</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>M</td>
<td>9</td>
<td>10</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>21</td>
<td>31</td>
<td>36</td>
<td>29</td>
</tr>
<tr>
<td>Nb rules</td>
<td>30</td>
<td>41</td>
<td>40</td>
<td>37</td>
<td>16</td>
</tr>
</tbody>
</table>
Corpus FoodTech En (3 millions tokens built from ISTEX Database)

<table>
<thead>
<tr>
<th>Rule</th>
<th>Nb</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>M-I-AN-N</td>
<td>A</td>
<td>R</td>
</tr>
<tr>
<td>M-I-EN-N</td>
<td>A</td>
<td>2821</td>
</tr>
<tr>
<td>M-I-NN-N</td>
<td>643</td>
<td>foodservice $\implies$ food information service</td>
</tr>
<tr>
<td>M-I2-(A</td>
<td>N)N-E</td>
<td>197</td>
</tr>
</tbody>
</table>
two words are in a semantic relation if they share the same lexical contexts. Many studies based on distributional paradigm [Hindle, 1990, Grefenstette, 1994, Lin, 1998, Hagiwara, 2008, Ferret, 2010]

**Context modelling**
- size of the context
- contextual items
- association measures

**Comparison of contexts**
- similarity measures
A clear distinction between synonyms and other semantically related words is not obvious [Lin et al., 2003, van der Plas and Tiedemann, 2006]

Semantic proximity:

- classical lexical relationships: synonymy, antonymy, hyperonymy, co-hyponymy, etc.
- non classical semantic relationships such as action/agent
Example of results of distributional approach for Fr, En and Es on [EOL]¹

<table>
<thead>
<tr>
<th>Rank</th>
<th>Fr: <em>ferme</em></th>
<th>En: <em>construction</em></th>
<th>Es: <em>torre</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>éolien</td>
<td>operation</td>
<td>aerogeneradores</td>
</tr>
<tr>
<td>2</td>
<td>horn</td>
<td>project</td>
<td>aerogenerador</td>
</tr>
<tr>
<td>3</td>
<td>aller</td>
<td>impact</td>
<td>pala</td>
</tr>
<tr>
<td>4</td>
<td>fréquence</td>
<td>road</td>
<td>rotor</td>
</tr>
<tr>
<td>5</td>
<td>puissance</td>
<td>require</td>
<td>estar</td>
</tr>
<tr>
<td>6</td>
<td>rev</td>
<td>aera</td>
<td>turbina</td>
</tr>
<tr>
<td>7</td>
<td>réseau</td>
<td>activity</td>
<td>altura</td>
</tr>
<tr>
<td>8</td>
<td>parc</td>
<td>plan</td>
<td>celosia</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>62</td>
<td></td>
<td>tower</td>
<td></td>
</tr>
</tbody>
</table>

¹corpus available at www.lina.univ-nantes.fr/?Ressources-linguistiques-du-projet.html
**Compositional method**

Assumption of semantic composition to generate the synonyms of a complex term

**First work [Hamon and Nazarenko, 2001]**

Synonymy of complex terms is compositional

\[
R_1 : T_1 = T_2 \land \text{syn}(E_1, E_2) \supset \text{syn}(CCT_1, CCT_2)
\]

\[
R_2 : E_1 = E_2 \land \text{syn}(T_1, T_2) \supset \text{syn}(CCT_1, CCT_2)
\]

\[
R_3 : \text{syn}(T_1, T_2) \land \text{syn}(E_1, E_2) \supset \text{syn}(CCT_1, CCT_2)
\]

With \( CCT_1 = (T_1, E_1) \); \( CCT_2 = (T_2, E_2) \) are complex terms and \( \text{syn}(CCT_1, CCT_2) \) a synonym relation between the candidate terms \( CT_1 \) and \( CT_2 \)

\( R_1 \) means that the heads are identical and the expansions are synonymous
A synonym of a MWT can be obtained by extracting synonyms and/or semantically related words of each component individually thanks to distributional methods.
**Synomymic MWT**

**Example**

- **Synonyms:** (EOL) En: *energy output/energy production* where *output* and *production* are synonyms (Termium)
- **Hyperonyms:** (EOL) Es: *implantación de las máquinas/implantación de aerogeneradores* where *máquina* is an hyperonym of *aerogenerador* (GDT)
- **Undefined:** (EOL) Fr: *arbre lent/arbre primaire* with no relation between *lent* and *primaire* (Terminalf)
Semi-Compositional Method

- Distributional Method
  - Provides synonyms of each lexical element of the MWT
  - Extend the rules from [Hamon and Nazarenko, 2001] to the semantic level and generalize them to MWT of any length
    - $R^G_1: T_1 = T_2 \land \text{sem}(E_1, E_2) \supset \text{sem}(CCT_1, CCT_2)$
    - $R^G_2: E_1 = E_2 \land \text{sem}(T_1, T_2) \supset \text{sem}(CCT_1, CCT_2)$
Examples of denominative variant discovered in [EOL] and [CAN]

### French

<table>
<thead>
<tr>
<th>N A → N A'</th>
<th>vents modérés</th>
<th>vents moyens</th>
</tr>
</thead>
<tbody>
<tr>
<td>A N → A' N</td>
<td>génératrice synchrone</td>
<td>alternateur synchrone</td>
</tr>
<tr>
<td>N P N → N P N'</td>
<td>coût de l’électricité</td>
<td>coût de l’énergie</td>
</tr>
<tr>
<td>N P N → N' P N</td>
<td>prix de l’électricité</td>
<td>coût de l’électricité</td>
</tr>
</tbody>
</table>

### English

<table>
<thead>
<tr>
<th>A N → A N'</th>
<th>initial surgery</th>
<th>initial operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A N → A' N</td>
<td>dynamic stall</td>
<td>static stall</td>
</tr>
<tr>
<td>N P N → N P N'</td>
<td>type of surgery</td>
<td>type of operation</td>
</tr>
<tr>
<td>N P N → N' P N</td>
<td>center of the blade</td>
<td>midpoint of the blade</td>
</tr>
<tr>
<td>N N → N N'</td>
<td>mastectomy swimsuit</td>
<td>mastectomy swimwear</td>
</tr>
<tr>
<td>N N → N' N</td>
<td>flow field</td>
<td>exchange field</td>
</tr>
</tbody>
</table>
Inferred variants
Patterns in sub-graphs

**Pattern**

\[ t_1 \rightarrow t_1' \]
\[ t_2 \rightarrow t_2' \]

- **VARIATION**
- **EXTENSION**

And if \( t_1' \) and \( t_2' \) share the same extension affix

**Example**

\[ n: \text{horizontal-axis} \]
\[ nnn: \text{horizontal-axis wind turbine} \]
\[ an: \text{horizontal axis} \]
\[ annn: \text{horizontal axis wind turbine} \]

- **VARIATION**
- **EXTENSION**

Same affix in common: "nn: wind turbine"

(Inferrred)
Inferred variants (2)
Mixing distributional and syntagmatic analysis

Morphological analysis

chemotherapy → hormone therapy
individual chemotherapy drugs → individual hormone therapy drugs

Syntagmatic analysis

(EOL) Fr: générateur synchrone → alternateur synchrone
(EOL) Fr: générateur synchrone à rotor → alternateur synchrone à rotor
TermSuite

https://github.com/termsuite

Documentation

http://termsuite.github.io/

TermSuite

Documentation

TerminologyExtractorCLI

- Usage
- Description
- Mandatory options
  - --from-text-corpus, --from-prepared-corpus
  - --tsv, --tbx, --json
- Other options
  - --capped-size INT
  - --context-associ-rate INT or FLOAT
  - --context-coccc-th INT or FLOAT
  - --context-scope INT
  - --contextualize (no arg)
  - --disable-derivative-splitting (no arg)
  - --disable-gathering (no arg)
  - --disable-merging (no arg)
  - --disable-morphology (no arg)
  - --disable-native-splitting (no arg)
  - --disable-post-processing (no arg)
  - --disable-prefix-splitting (no arg)
  - --enable-semantic-gathering (no arg)
  - --encoding, -e ENC
  - --from-prepared-corpus DIR
## Output of TermSuite with TSV format

Extract from [EOL]

<table>
<thead>
<tr>
<th>type</th>
<th>pilot</th>
<th>freq</th>
<th>spec</th>
<th>semScore</th>
<th>isDico</th>
<th>isDistrib</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>T rotor</td>
<td>848</td>
<td>4.82</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>T wind turbine</td>
<td>1855</td>
<td>4.56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>V[h]+ wind power-plant</td>
<td>2</td>
<td>1.90</td>
<td>0.97</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>V[h]+ wind channel</td>
<td>2</td>
<td>2.20</td>
<td>0.97</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>V[h]+ Wind turbines-a</td>
<td>2</td>
<td>2.20</td>
<td>0.97</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>V[h]+ wind farm</td>
<td>488</td>
<td>3.20</td>
<td>0.89</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>V[s] wind turbine rotor</td>
<td>31</td>
<td>3.38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>V[s]+ vertical-axis wind turbine</td>
<td>6</td>
<td>2.37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>V[s] WIND TURBINE APPLICATIONS</td>
<td>86</td>
<td>3.83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>V[s] wind turbine blades</td>
<td>48</td>
<td>3.57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>V[h]+ Enfield-Andreau turbine</td>
<td>3</td>
<td>2.37</td>
<td>0.54</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>V[s]+ wind turbine concepts</td>
<td>37</td>
<td>3.46</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>V[s]+ wind turbine generator</td>
<td>27</td>
<td>3.02</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>V[s]+ Domestic Wind Turbines</td>
<td>29</td>
<td>3.35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>V[s]+ small wind turbines</td>
<td>33</td>
<td>3.41</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>V[s] MW wind turbine</td>
<td>10</td>
<td>2.89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>T wind power</td>
<td>278</td>
<td>4.34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>V[s] wind turbine power</td>
<td>10</td>
<td>2.89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>V[s]+ wind power stations</td>
<td>24</td>
<td>3.27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>V[s] Wind Power Project</td>
<td>19</td>
<td>3.17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>V[s] wind power development</td>
<td>14</td>
<td>3.04</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>V[s]+ wind power generation</td>
<td>11</td>
<td>2.63</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>V[s]+ wind power capacity</td>
<td>6</td>
<td>2.67</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>V[s]+ wind power penetration</td>
<td>4</td>
<td>2.50</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>V[s] Wind Power Installation</td>
<td>2</td>
<td>2.20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>T airfoil</td>
<td>236</td>
<td>4.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>T voltage</td>
<td>214</td>
<td>4.22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>T TURBINE SOUND</td>
<td>80</td>
<td>3.80</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>V[s]+ WIND TURBINE SOUND</td>
<td>71</td>
<td>3.74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>V[h]+ turbine noise</td>
<td>52</td>
<td>3.61</td>
<td>0.65</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
TermSuite users

TermSuite was initiated during the research project TTC (*/*FP7/2007-2013*/*) under Grant Agreement no. 248005 from 2010 to 2012 and has been supported by ISTEX, French Excellence Initiative of Scientific and Technical Information, from 2015 to 2017

INIST: book indexation
Meteojob: alignment cv and job advertisements
Health oriented Innovation Lab (US): clinical NLP
University of the Basque country: specific bilingual dictionaries for MT
Prometil: detect errors in your requirements
Conclusion and Future works

Discovering terms and variants in a multilingual framework using both syntagmatic and distributional analysis
Conclusion and Future works

Discovering terms and variants in a multilingual framework using both syntagmatic and distributional analysis

**Perspectives**
- Verbal variants
- Distributional analysis
Conclusion and Future works

Discovering terms and variants in a multilingual framework using both syntagmatic and distributional analysis

**Perspectives**

- Verbal variants
- Distributional analysis
  - paraphrases
Conclusion and Future works

Discovering terms and variants in a multilingual framework using both syntagmatic and distributional analysis

Perspectives

- Verbal variants
- Distributional analysis
  - paraphrases
  - word embeddings
Discovering terms and variants in a multilingual framework using both syntagmatic and distributional analysis

**Perspectives**

- Verbal variants
- Distributional analysis
  - paraphrases
  - word embeddings
- Multilingual variants
References


*Term variation in specialised corpora characterisation, automatic discovery and applications.*

B. Daille

Collection

Terminology and Lexicography Research and Practice

John Benjamins (printing)
Thank you for your attention
Thank you for your attention

QUESTIONS?


Is it possible to predetermine a referent included in a French noun structure?


Automatic retrieval and clustering of similar words.

Identifying synonyms among distributionally similar words.

Finding synonyms using automatic word alignment and measures of distributional similarity.
In 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics ACL’06, Sydney, Australia.