TreeCloud & Unitex: an increased synergy
Claude Martineau

To cite this version:
Claude Martineau. TreeCloud & Unitex: an increased synergy. ECLAVIT Workshop, Nov 2017, Champs-sur-Marne, France. 2017. hal-01702091

HAL Id: hal-01702091
https://hal-upec-upem.archives-ouvertes.fr/hal-01702091
Submitted on 19 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
TreeCloud & Unitex: an increased synergy
Claude Martineau
Projet PEPS CNRS/UPE Eclavit

TreeCloud is a tree cloud visualization of a text

TreeCloud builds a tree cloud visualization of a text, which looks like a tag cloud where the tags are displayed around a tree to reflect the co-occurrence distance between the words in the text.

The adjacent grammar contains three path/bases:
• The grey box represents a call to a subgraph that recognizes auxiliary or modal verbs.
• The second box uses some lexical masks (e.g., DICO1, DICO2) to match grammatical category.
• The 'ANTIPERFOS-basal mask in the last box matches a list of stopwords given by the user. Each word of this list is labelled/tagged with ANTIPERFOS-
  A node with numbers.

Some examples of trees in different languages

The 45 most frequent compound and person nouns in a Latin Verde novel

The 30 most frequent words in a Serbian article

Conclusion
Plug-in of Unitex into TreeCloud provides:
• A more accurate representation of forbidden words
• All kinds of modalities to be recognized in the tree
• A visual representation of some grammatical or semantic categories of the words.
• A faster construction of the tree (via a careful use of the Unitex API)

Further details of the TreeCloud/Unitex integration are available at:
http://treecloud.univ-lr.fr/