

# XIP German Grammar

(september 2006, S. Maurel)

## 1. Introduction

The purpose of this document is to list the existing files that are needed by the XIP German grammar (by august 2006).

## 2. List of files

**german.xip, german.grm** (in the repertory GRMFILES)  
declare the basics.

**disamb\_agr.xip** (in DISAMB),  
**disamb\_lemma.xip** (in DISAMB),  
**disamb\_rank.xip** (in DISAMB),  
**disamb\_tagger.xip** (in DISAMB),  
**chunker.xip** (in CHUNKING),  
**dependencies.xip** (in DEPENDENCIES),  
contain the core modules, their functions are detailed in the documents of the repertory DOC.  
(NB: there are files with the suffix **\_PE**, they are for the application Pocket Engineer, they are not necessary for the general German grammar!)

**categories\_tagger.xip** and **categories.xip** are for the categories of the chunks,  
**controls\_tagger.xip** and **controls.xip** contain some control values,  
**features\_tagger.xip** and **features.xip** define the features,  
all of these files are in FEATURES.

**functions.xip** (in DEPENDENCIES)  
contains the names of the dependencies.

**translation\_tagger.xip** (in FEATURES)  
contains the features to convert the tags from the tagger to XIP.

**lexicon\_tagger.xip** and **lexicon.xip** (in LEXICONS)  
some words or features of words which are not in the FST can be added in this file.

**localgrams.xip** (in LOCALGRAMS)  
contains multiple words, product names, etc. which are not desired in the FST because they are too specific.

The corpus files (different articles from newspapers and some law texts) are in the repertory CORPUS.

The repertory EVAL contains the text files which were used to do the evaluation, for each text there is the original text file, the XIP result file and the manually annotated result file. All results are counted in an excel table and resumed in a word document.