Multi-Lingual Noun Phrase Extractor (MuNPEx) v1.0 for GATE released
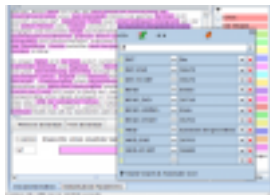
- GATE Components
- MuNPEx
- Noun Phrase Chunking

MuNPEx 1.0 The noun phrase chunker MuNPEx *(Multi-Lingual Noun Phrase Extractor)* is now available in the new and improved release v1.0. MuNPEx is a base NP chunker for the GATE framework and implemented in JAPE. It is fast, robust, customizable, well-tested and currently supports English, German, and French (with Spanish in beta).

Major changes in this release:

- Limited number of pre- and post-head modifiers to make MuNPEx more robust on certain kinds of input (like a long list of tags or menu entries when processing web pages)
- New optional grammars to add a HEAD_LEMMA slot to an NP annotation, with the lemma extracted from the GATE morphological analyser (for English), the Durm Lemmatizer (for German), or the TreeTagger (for German, Spanish, French)
- DET/MOD/HEAD/MOD2 slots are now stored as strings (rather than Content objects) to make them easier to export and compatible with the new Predicate-Argument Extractor (PAX) component
- other code cleanup and improvements
- no longer labeled as "beta" -- five years of testing ought to be enough, we're not Google ;-)

For more details and the download, please visit the MuNPEx page.

Source URL (retrieved on *2025-12-22 21:10*): https://www.semanticsoftware.info/noun-phrase-chunker-munpex-release-1-0