

How to create a MWE lexical entry

Aleksandar Petrovski, Katerina Zdravkova

Working group 1

Classification presented by Baldwin and Kim has been used. MWEs are classified by three main properties: flexibility, idiomaticity and syntactic structure.

Property	Value	Code
Fixedness/Flexibility	Fixed	FxF
	Semi-fixed	SxF
	Flexible	FlxF
Idiomaticity	Lexical	LexI
	Syntactic	SynI
	Semantic	SemI
	Pragmatic	PrgI
	Statistical	Stal
Syntactic structure	Nominal	NomS
	Verbal	VerS
	Adjectival	AdjS
	Adverbial	AdvS

How to create a MWE lexical entry

кабелска телевизија, N+FLX=N2 +SFxF+PrgI+StaI+NomS

NooJ formalism: *кабелска телевизија* is a MWE (eng. cable television), *N* stands for noun - grammatical category of the MWE, *N2* is the inflectional class code that describes the inflectional paradigm, *SFxF*, *PrgI*, *StaI* and *NomS* are codes from Table 1, describing MWE's properties.

Since *N2* describes the lexical entry's inflectional behaviour, NooJ is able to recognise all of its 9 inflectional forms.

Note: commas are not allowed within MWEs.

- NooJ lexicons are very good when dealing with fixed and semi-flexed MWEs. For flexible MWEs, due to possible change of word order, lexicons can't help much. The only way is to build a syntactical grammar for each MWE.
- It would be interesting to analyse how other NLP tools deal with flexible MWEs and to compare them with NooJ.