

## overview

**implementation:** paper-and-pencil analyses → electronic resource  
**challenge of MWEs:** capturing idiosyncrasies and regularities in the interplay of morphology, syntax and semantics

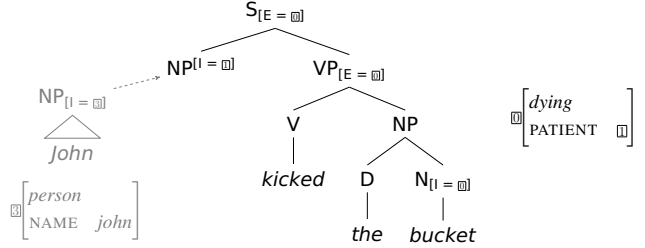
**How to allow for an adequate and user-friendly implementation of MWEs?**

**technical requirements:**

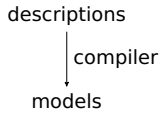
- support of linguistic description levels and their interaction
  - support of different description languages and models
  - largely theory neutral
  - easy access
- ⇒ **eXtensible MetaGrammar (XMG)**

## MWEs in LTAG

classic approach causing syntactic ambiguity:<sup>[1,2]</sup>



## implementation with XMG



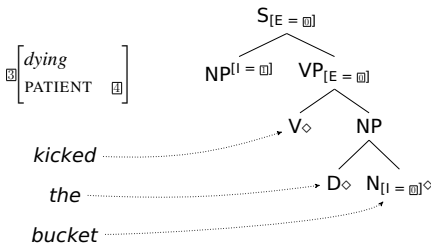
**eXtensible MetaGrammar<sup>[3]</sup>**

- XMG descriptions are declarative and multidimensional.
- XMG descriptions are packed into classes.
- XMG classes make up an inheritance hierarchy.

**dimensions with specific description languages**

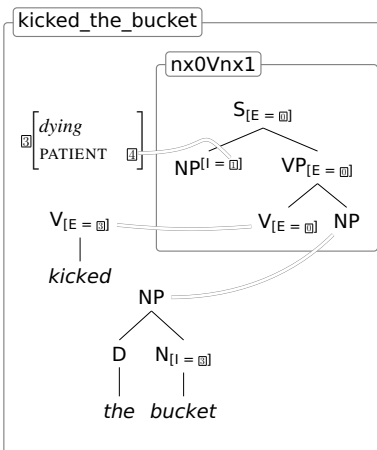
- <syn>: tree descriptions
- <sem>: underspecified formulae of predicate logic
- <frame>: TFS descriptions → semantic frames<sup>[5]</sup>
- <morph>: morphological descriptions<sup>[4,6]</sup>

## implementation with external anchoring



- **Standard LTAG approach:** dissociate the **lexical anchors** from the **tree template**.
- The tree template is described by the metagrammar, whereas the lexical anchors are dealt with as part of a two-level lexicon.
- In the two-level lexicon, full forms are mapped onto lemmas, and lemmas are again mapped onto tree templates or tree families.
- **Problem of this approach:** does not easily allow for attaching the lexical semantics in such way that the intended linking is established.

## implementation with internal anchoring



```
class nx0Vnx1
export ?S ?NP_Subj ?VP ?V ?NP_Obj
declare ?S ?NP_Subj ?VP ?V ?NP_Obj ?X0 ?X1
{
  <syn>{
    ?S [cat=s, e=?X0] {
      ?NP_Subj [cat=np, i=?X1]
      ?VP [cat=vp, e=?X0] {
        ?V [cat=v, e=?X0]
        ?NP_Obj [cat=np] }}
  }
}
```



```
class kicked_the_bucket
import nx0Vnx1[]
declare ?X3 ?X4
{
  <syn>{
    ?NP_Subj [i=?X4];
    ?V [e=?X3] "kicked";
    ?NP_Obj [] {
      [cat=det] "the"
      [cat=n, i=?X3] "bucket" }
  }
  <frame>{
    ?X3[dying,
      patient:?X4]
  }
}
```

- The metagrammar describes complete elementary trees and their semantics.
- nx0Vnx1 contributes the generic syntactic structure of a transitive verb.
- kicked\_the\_bucket reuses nx0Vnx1 and adds to it the lexical anchors and their idiomatic semantics.
- The sharing of variables across <frame> and <syn> is responsible for the linking between syntactic positions and positions in the semantics.

[1] Abellé, A. 1995. The flexibility of French idioms: a representation with lexicalized tree adjoining grammar. In E.-J. Everaert M. and van der Linden, A. Schenk & R. Schreuder (eds.), *Idioms: structural and psychological perspectives*, 15-42. Lawrence Erlbaum.  
[2] Abellé, A. & Y. Schabes. 1996. Non-compositional discontinuous constituents in tree adjoining grammar. In H. Bunt & A. van Horck (eds.), *Discontinuous constituency*, 279-306. Berlin: Mouton de Gruyter. [3] Crabbé, B., D. Duchier, C. Gardent, J. Le Roux & Y. Parmentier. 2013. XMG: eXtensible MetaGrammar. *Computational Linguistics* 39(3). 1-66. [4] Duchier, D., B. M. Ekoukou, Y. Parmentier, S. Petitjean & E. Schang. 2012. Describing morphologically rich languages using metagrammars: A look at verbs in Ikota. In *Workshop on language technology for normalisation of less-resourced languages (SALTMIL 8 - ALLaT 2012)*, 55-59. [5] Lichte, T. & S. Petitjean. 2015. Implementing semantic frames as typed feature structures with XMG. *Journal of Language Modelling* 3(1). 185-228. [6] Petitjean, S., Y. Sami & T. Lichte. 2015. Une métagrammaire de l'interface morpho-sémantique dans les verbes en arabe. In *Actes de la 22e conférence sur le traitement automatique des langues naturelles*, 473-479. Caen, France.