Typological databases, lexical databases, and corpus resources

The databases and resources discussed here can be accessed from the Resources page on the course web site; go to "Morphological data from natural languages" and follow the links found there.

(1)   Database focused on specific categories of linguistic/typological information

    (a)   What are some examples from the Surrey Morphology Group collection?

    (b)   What is the interface typically like? / What functions are often supplied?

    (c)   What kinds of research is it good for?

(2)   Lexical database

- Try: MorphIt (read description)
  MRC Psycholinguistic Database

    (a)   What does it typically consist of?
  - lexemes, wordforms, paradigms
  - frequency information
  - relationships among lexemes (semantic, morphological, etc.)

    (b)   What is the interface typically like? / What functions are often supplied?

    (c)   What kinds of research is it good for?

(3)   Corpus

- Try: Querying Internet corpora

    (a)   What does it typically consist of?
  - collection of 'texts' (written or spoken)
  - may be *parsed* (annotated for morphological and/or syntactic information)

    (b)   What is the interface typically like? / What functions are often supplied?
  - concordance
  - collocation

    (c)   What kinds of research is it good for?