ANN'S OPUS

- definition of MWE: want to avoid committing to as specific definition as "includes at least one space" (question of what are segments as far as the parser is concerned, e.g. LKB treats hyphens as spaces; also question of variability in hyphenation, and when to use and when not)

- sort out what is specific to MWEs from what is not (lighten the load)

- lexicalised vs. institutionalised MWEs: *institutionalised phrases are syntactically and semantically predictable but are used unexpectedly frequently*

DELIMITING MWEs

- *brief moment*: fully compositional but semantically redundant (tautology) → MWE?

- *foreign direct investment*: fully compositional in- and out-of-domain, but idiosyncratic adjective ordering → MWE?

- *at/*in/*on nine o'clock*: fully compositional, but idiosyncratic preposition selection → MWE?

SYNTACTIC MARKEDNESS

- What constitutes syntactic markedness? We can test whether a given expression undergoes typical syntactic variation, but at what point can we say that the expression is syntactically marked?

  - article/determiner variation (N-N, Adj-N)

  - N1-N2/N2-P-N1 alternation (N-N)

  - number (N-N, Adj-N)

  - voice, mood (phrasal verbs)

  - object insertion (phrasal verbs)

  - modifiability/quantifiability

  - topicalisability/ellidability of certain parts of the expression (verbal expressions)

  - pronominalisability of certain parts of the expression

- passivisability (verbal expressions)

- adverb insertability (verbal expressions)

- coordination

- relativisability (adjectival/verbal expressions)

- attributive/predicative alternation (adjective–noun constructions)

- part-wise semantic correlation with other MWEs

- *the universe*: fully compositional but always takes the definite article → MWE? If not what and why?

- I.e. is every phrase that is marginally idiosyncratic a MWE? If not, what combination of features demarcates MWEs from productive phrases?

ANTI-COLLOCATIONS

- Problems with anti-collocations: what are they and where do they come from? I.e. are they determined by:

  - (synonym) substitution from the canonical form (e.g. *red/*blue herring/#mackerel/#fish*)

  - or from grammatical processes (e.g. *man of letters* vs. **letters man*)

  - or are they alternative lexicalisations of the base concept (e.g. *rolling blackouts* vs. *sequenced outages*)