

Using TIGERSearch
L615
Spring 2009

Getting started On one of the computers in the lab in MM401:

1. Open up a terminal
2. Type something like the following (it may differ depending on exactly where TIGERSearch is installed): `cd /Applications/TIGERSearch/lib/`
3. Type `./runTS &`
 - This should open up TIGERSearch

Finding your way around

- We have a few syntactically-annotated corpora installed: for tonight, we'll mostly use the Penn Treebank (WSJ), but note how you can load other corpora
- We'll use the graphical interface, but note how you can generate a textual equivalent
- For individual tokens, you can search for words and/or POS forms, including things like “starts with” and some negation
- For nonterminal categories, you can do similar things
- The searchable *vertical* relations between nodes are: direct dominance & dominance (transitive closure)
- The searchable *horizontal* relations between nodes are: direct linear precedence & linear precedence (transitive closure)

Searches to try

1. What are the different argument structures for the verbal *need* and with what frequency?
2. What (non-modal) verbs take other verbs as arguments? What's the longest sequence of embedded verbs (e.g., *might have been going*)? (NB: syntactic structure is potentially more reliable than POS sequences here, given the possibility of intervening adverbs & parentheticals)
3. What verbs license inversion sentences (e.g., *Into the room walked John*)?
4. How often do noun-modifying relative clauses use *that?* *which?* neither (e.g., *friends I know*)?
5. When can comparative correlatives (e.g., *the faster the car (is), the better (it is)*) have verb deletion?

Or search for whatever interests you ...