THE FEL-OU M CORPUS PROJECT

Final report

Hazidi Abdul Hamid
What was our objective?

- To gain an understanding of the mastery of English at the syntactical level among the mature ODL learners based on corpus analysis.
What did we study?

- We studied a corpus that we built ourselves
The corpus we built

- >29K assignment papers
  - OUMH2103 & OUMH2203.
    - 100% in English
    - 100% submitted online
    - Taken by all students (all subjects)

- >15 Million words
The random sampling

- 5000 papers selected
  - Random selection
  - Cleaned for bugs
The corpus analysis tool

- Wordsmith 6.0 - [http://www.lexically.net/wordsmith/](http://www.lexically.net/wordsmith/)

- What we looked for:
  - **Collocates**: words that occur together. We only used,
    - **R1**: target word & first word to the right. E.g.: BE + given
    - **L1**: target word & first word to the left. E.g.: Must + BE
  - **Clusters**: cluster of words that form the environment with the +R1 or +L1 in the middle: cluster of 3 words to the left and right.
What we looked for?

- Selected members of the closed word classes: prepositions, auxiliary verbs, modals, etc.
  - *BE*: *am, is, are,*
  - *HAVE*: *has, had,*
  - *CAN*
Why did we look for these?

- They do not change
- They are essential elements of core grammatical constituents
What we found?

- Analysis did not show significant ungrammatical construction within the R1 & L1 pairs.
- Confirmed by analysis of the clusters.
Where there ungrammatical constructions?

- Yes but they are at the prosodic level and beyond
What does this imply?

- The subjects know how to build the actual phrases (NP, VP, AdvP, PP, etc.)
- The mistakes that the subjects did were at the prosodic level and beyond: idiomatics, semantics, pragmatics, conventions etc.
How does this affect OUM?

- Our students can process formal grammatical content in re-acquaintance / revision courses.
- It may be prudent to teach them English in context: English for Specific Purposes.