A corpus-based approach to LSP: using Intellitext to create study materials for Dentistry 1

Marion Bowman (School of Dentistry, UoLeeds)

&

James Wilson (School of Modern Languages and Cultures, UoLeeds)

UNIVERSITY OF LEEDS

START

KUWAIT

+

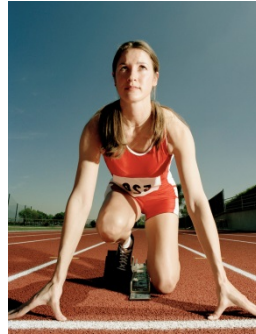11-13 places: IFY   4 places

- Total cohort of approx. 90
- Tough 5 year course
- Performance gap

Initial challenges

- Type of **provision** for international students (equity)
- Lack of **integration** between international and home students
- How to identify **key target language** and study strategies for Dental Surgery 1
- **Non-specialist** (language teacher/academic adviser – some awareness of corpus-based teaching)
- **Crammed timetable** and time-poor students

Collaboration

James Wilson:
-corpus-based
learning
-Intellitext

Marion Bowman:
- curriculum,
- context,
- assignments

Language at Leeds project

Design

- 'Quick and dirty corpus' (Chris Tribble) (Adv/Disadv)
- Data input (summer 2013)
- Lecture notes, e-lecture commentary, key online texts
- 4 modules (sub-corpora) and 1 Dentistry corpus
- Materials design for 8 Language of Dentistry (LOD) taught small group sessions
- 1 session on how to access database

UNIVERSITY OF LEEDS

- Numbers attending optional 1 hr weekly sessions (Sem1):

	comm	hhp	hhp	hhp	ioe	intell	ioe	ioe
TTL	19	16	8	10	13	13	10	0
INT	5	4	0	0	2	4	2	0

- NS, NNS, ESL, dyslexia – complex mix
- 'Keen bean' effect?
- Have a look at some example materials from the task- based sessions (backwards design; technical vocab & language in between)…

IntelliText

- IntelliText (http://smlc09.leeds.ac.uk/itb) …
  - allows access to monolingual corpora for 12 languages (there is more than one corpus for most languages)
  - allows access to several bilingual corpora
  - has 8 search functions
  - can be used generate keyword and frequency lists
  - includes a "Build Your Own Corpus" function that allows users to create and annotate their own corpora
  - is freely available: it can be used by anybody anywhere

- The Dentistry 1 corpus was created from lecture notes, mainly PowerPoint slides, used on the Dentistry 1 module

- The corpus contains approximately 200,000 words

- There are four subcorpora that can uploaded and analysed separately

- Keyword lists (of single and multi-words) were extracted and used as the basis for materials design

- The lists can be saved as Excel files and cleaned manually

- Keywords were entered into IntelliText concordance and collocation searches

Sample searches (keywords)

UNIVERSITY OF LEEDS

Key Collocations

Collocation	Count	F1	F2	LL	MI	T
tooth surface	50	694	347	84.71	6.1	6.97
tooth morphology	14	694	35	31.53	7.57	3.72
tooth support	7	694	80	9.81	5.38	2.58
tooth decay	3	694	6	7.17	7.89	1.72
tooth crown	5	694	54	7.15	5.46	2.19
tooth type	6	694	129	6.51	4.46	2.34
tooth loss	5	694	139	4.8	4.09	2.11
tooth cleanliness	2	694	5	4.49	7.57	1.41
tooth alignment	2	694	6	4.26	7.31	1.41
tooth eruption	2	694	7	4.08	7.08	1.4
tooth development	3	694	38	4.04	5.23	1.69
tooth site	4	694	143	3.36	3.73	1.85
tooth relationship	3	694	60	3.35	4.57	1.66
tooth chronology	1	694	1	3.08	8.89	1

Dentistry 1 Keywords

Lemma	F1	F2	LL
periodontal	85	434	5219
tooth	5263	694	4813
dental	2666	403	2897
patient	41031	689	2117
plaque	993	225	1783
oral	4911	278	1486
scaler	25	122	1463
gingival	9	106	1325
periodontitis	9	99	1234
bone	6509	262	1231
surface	16342	347	1215
tissue	5267	235	1150
probe	2993	184	1012
disease	29958	370	934
enamel	344	109	928

Extended text

Concordances

titleid	left	match	right
>>	imperfecta (Poorly formed tooth	enamel	) N. B. Lyonisation – resulting
>>	produce even more acid, so that	enamel	eventually demineralizes) Endogenous
>>	the tooth that is visible is the	enamel	of the tooth crown. Some of the
>>	of the tooth crown. Some of the	enamel	is hidden beneath the gum. Enamel
>>	enamel is hidden beneath the gum.	Enamel	is the white, hard wearing outer
>>	clinical ' crown. As some of the user	enamel	is normally obscured by the gum
>>	can see the whole extent of the	enamel	. By definition, the part of the
>>	the part of the tooth covered in	enamel	is called the ' anatomical ' crown
>>	larger than the clinical crown.:	Enamel	is a very brittle and hard- wearing
>>	Indeed a tooth made totally of	enamel	would be very resistant to wear
>>	Much of the microstructure of	enamel	has evolved to try and reduce
>>	possible lines of shear within the	enamel	itself to prevent fractures propagating
>>	propagating within it. Although	enamel	forms the outer covering of the
>>	the dentine and also because the	enamel	is gradually wearing thinner so
>>	dentine shows through more clearly.	Enamel	does not have a cell population

Materials design

- Exercises were created (1) to encourage active engagement with IntelliText ("hands-on approach") and (2) from material in the corpus to test students on near synonyms ("hands-off approach")

- Our focus was on both discipline-specific terminology and on general language

- International students benefited from seeing the dentistry-specific terms in context

- Home students found the discipline-specific keyword lists useful for their revision

UNIVERSITY OF LEEDS

"give" vs. "administer" vs. "deliver"

	give	administer	deliver
advice			
supplemental oxygen			
tangible results			
postoperative instructions			
consent			
the correct dose of			
an indication of			
drugs			

UNIVERSITY OF LEEDS

Compare:

- The driver **delivered** the drugs to the hospital.
- Nurses **administer** the drugs every four hours.
- The doctor **gave** him drugs to combat the pain.

- Students are encouraged to search for other examples in the Dentistry 1 corpus as well as reference corpora to highlight trends in the use of similar words

UNIVERSITY OF LEEDS

NS: Why did you choose to attend LOD sessions?

- Revision (4)
- Ease transition to uni (3)
- Curiosity / seemed helpful (3)
- Understand the course structure (1)
- Language / terminology (1)
- Dyslexia (1)
- Referral (1)

Evaluation of usefulness

NS: Why did you find sessions useful?: Revision, Transition, Content knowledge, Enhancement, Writing.

- "Relaxed, enjoyable, helped me to structure my learning and recap – communication as revision".

NS: Why did you find sessions not useful?

- "Some parts were only useful for people whose first language is not English".

- 50% of NNS respondents used the Manchester phrasebank v 20% of NS respondents

- 0% of NNS used the Intellitext database, 10% of NS used in on own.

Conclusions

- Corpus allowed non-specialist to develop accurately targeted materials for workshop design
- Corpus is flexible enough for future development (transferrable approach)
- Sessions evaluated well (NS and NNS)
- NS appreciated study strategies and content vocab (useful for revision)
- NNS attendance dropped off
- NNS (poor response on questionnaire)
- Low rate of self-access of corpus (NS and NNS)

- Put one session in timetable for all year 1's
- Get permission to target NNS only
- Optional but with pressure
- Mini-corpus for postgrads?
- Corpora for curriculum strands?