

# Language & IT

## Dr. Abdullah Al Fraidan

### Lecture 11-BNC and suggested tasks

#### Corpus Linguistics

What is BNC?

<http://www.natcorp.ox.ac.uk/>

#### المهام المقترحة-Suggested Tasks

##### SOME SUGGESTED CORPUS TASKS

- بعض المهام CORPUS المقترحة
- The important thing to realise of course is that corpora and search engines primarily constitute tools or research methods, rather than areas of enquiry in themselves.
- الشيء المهم أن ندرك بالطبع هو أن corpora ومحركات البحث تشكل أساساً أدوات أو مناهج البحث، بدلاً من مجالات التحقيق في حد ذاتها.
- Few people study hammers; rather they use them to perform tasks which they think of, like building a cupboard, and which have dimensions remote from hammers which the user has to bring a lot of separate expertise to.
- عدد قليل من الناس لدراسة المطارق ، استخدموا لتنفيذ المهام التي يفكرون، مثل بناء خزانة، والتي لها أبعاد بعيدة عن المطارق التي كان المستخدم لتحقيق الكثير من الخبرة المنفصلة.
- Similarly, corpus use, like introspection or administering tests or questionnaires to subjects, is not in itself usually a project in itself.
- وبالمثل، استخدام كوربوس، مثل التأمل أو إجراء اختبارات أو استبيانات لمواضيع، ليس في حد ذاته عادة ما يكون المشروع في حد ذاته
- Rather it is a means to carry out some project in language description, language teaching or whatever. .
- بل هو وسيلة لتنفيذ بعض المشاريع في وصف اللغة، تعليم اللغة أو أي كان. .

- The bulk of the project has to come from the user's prior knowledge of linguistics, teaching etc.
- الجزء الأكبر من المشروع يجب أن يأتي من المعرفة المسبقة لمستخدم اللغويات، وتعليم الخ
- Therefore in choosing a task you have to think what linguistics you know most about already, and choose a task accordingly.
- لذلك في اختيار المهمة التي يجب أن نفكر ما اللغويات انك تعرف أكثر حول بالفعل ، واختيار المهمة وفقا لذلك.
- Those suggested below are mostly descriptive linguistic (vocabulary and grammar mainly), or involve some pedagogical evaluation or authoring with a corpus element.
- تلك المقترحة أدناه هي في معظمها لغوية الوصف (المفردات والقواعد في المقام الأول)، أو تنطوي على بعض التقييم التربوي أو التأليف مع وجود عنصر الإحضار.
- They should be do-able with the corpus and concordancing resources you can access this year, though I cannot guarantee anything as what are available changes by the minute.
- ينبغي أن تكون قادرة بالفعل مع كوربوس وفهرسة الموارد يمكن الوصول لهذا العام، بالرغم من أنني لا يمكن أن أضمن أي شيء مثل ما هي التغييرات المتاحة قبل لحظة.
- You can of course also think of your own projects in accordance with your own interest, and to connect with other courses you may be doing (since almost any course you take in the Department of Language and Linguistics potentially has a corpus dimension).
- أيضا يمكنك بالطبع تفكر في المشاريع الخاصة بك وفقا للمصلحة الخاصة بك، وللتواصل مع غيرها من الدورات التي قد تكون به (منذ ما يقرب من أي بالطبع كنت تأخذ في قسم اللغة واللغويات المحتملة لديه لأبعاد كوربوس).
- The following are not fully worked out, and in no particular order. You have to decide if the relevant corpus is available, and of course get the search engine to dig out useful information.
- لا يتم عمل ما يلي بشكل كامل، ودون ترتيب معين. عليك أن تقرر ما إذا كان كوربوس ذات صلة متاحة، وبطبيعة الحال الحصول على محرك بحث للخروج بمعلومات مفيدة.
- Often the instructions you can enter in the search engine will not produce all and only the information you want.
- غالبا فإن التعليمات التي يمكن أن تدخل في محرك البحث لا تنتج الكل فقط المعلومات التي تريدها

- The trick is to get as close as possible and then sort through the output by hand for what you need..... and interpret it.
- هو خدعة للحصول في أقرب وقت ممكن ومن ثم من خلال فرز الناتج باليد لما تحتاج ..... وتفسير ذلك.
- Make good use of your intuitions as a teacher and/or descriptive linguist!
- الاستفادة من حدسك كمدرس و / أو واصف لغوي!

### - المهام المقترحة - Suggested Tasks

- How do 'synonyms' differ?  
كيف 'المرادفات' تختلف؟
- In class we look briefly at the 'synonyms' sad and unhappy.  
في الفئة ونحن ننظر بايجاز في حزين 'المرادفات' وغير السعيدة
- You could look at another pair of 'synonyms' like any of those below.  
هل يمكن أن ننظر إلى زوج آخر من "المرادفات" مثل أي من تلك أدناه.
- Read about synonymy in Ullmann, Leech, Zgusta etc. so as to have in mind the different KINDS of ways in which they may subtly differ.  
اقرأ عن الترادف في أولمان، ليبش، Zgusta الخ وذلك ليكون في اعتبارها أنواع مختلفة من الطرق التي قد تختلف بمهارة.
- Get concordance output from a suitable set of texts.  
الحصول على إخراج التوافق من مجموعة مناسبة من النصوص.
- Give an account of the similarities and differences between the chosen synonyms.  
اعط حساب وأوجه الشبه والاختلاف بين المرادفات المختارة.
- How much of your analysis is from the corpus, how much from introspection prompted by the corpus information?  
وكم من التحليل الخاص بك من كوربوس، وكم من التأمل المطالبة من قبل المعلومات كوربوس؟
- You could also refer to entries in dictionaries of synonyms which don't just list them but include 'synonym essays' attempting to explain the differences (e.g. Webster's Dictionary of Synonyms, or Cassell's Modern Guide to Synonyms which is available in Colchester Public Library, Trinity Square, town centre) both as a source of ideas as to what the differences are and something to criticise.  
هل يمكن أن تشير أيضا إلى الإدخالات في القواميس من المرادفات التي لا مجرد سرد لهم ولكن تشمل 'مرادف المقالات' محاولة لشرح الفروق (مثلا قاموس وبستر من المترادفات، أو الدليل الحديث للمرادفات والذي يتوفر في كولشستر المكتبة العامة، ساحة المركز الثالث كاسيل، البلدة) على حد سواء باعتبارها مصدرا للأفكار على ما هي أختلافات وشيء للأنقاد.

- See also error and usage books like Alexander: Right Word Wrong Word and Heaton and Turton: Dictionary of Common Errors.
- أيضا انظر خطأ واستخدام الكتب مثل الكسندر: كلمة صحيحة كلمة خاطئة وهيتون وTURTON: معجم الأخطاء الشائعة
- These are for foreign learners.
- هذه هي لمتعلمين الأجنبية
- You could use them as an aid to your own analysis or do a critique - do they mislead?
- هل يمكن استخدامها كوسيلة مساعدة لتحليلك أو القيام بالانتقاد - هل التضليل؟
- If you are interested in language teaching you might like to think how you might select and adapt the corpus lines you find to make an effective synonym differentiation task for some specific learners you have in mind.
- إذا كنت مهتما في تدريس اللغة قد ترغب في التفكير كيف يمكن تحديد وتكييف خطوط كوربوس لتجد مهمة فعالة لتفريق المرادف لبعض المتعلمين محددة لديك في الاعتبار.
- You could also consider synonyms within some specific variety of English (e.g. academic writing) rather than overall, by choosing a corpus within the BNC, for example.
- لك ان تنتظر أيضا في بعض المرادفات المتنوعة المحددة من الإنجليزية (مثل الكتابة الأكاديمية) بدلا من العام، عن طريق اختيار كوربوس داخل BNC، مثلا.

### Suggested Tasks

- Distinguishing 'confusibles'.
- المميّزة
- There are some well-known 'confusibles' which are similar in sound as well as meaning (cf dictionaries by Room).
- هناك بعض 'confusibles' المعروفة التي تشبه في الصوت وكذلك المعنى (راجع قواميس الغرفة).
- Some of these may also have proscriptions associated with them. See also the points made about the last task.
- بعض هذه قد يكون لها أيضا محظورات مرتبطة بها. انظر أيضا النقاط التي وردت حول المهمة الأخيرة.

- Corpus research should help one to sort out what really are the differences, and whether statements in dictionaries (e.g. usage notes in Longman Dictionary of Contemporary English or the Oxford Advanced Learners Dictionary) about how they are used are correct.
- يجب لبحوث كوربوس مساعدة احد لفرز ما هي حقا الاختلافات، وعما إذا كانت تصريحات في القواميس (مثلا تستخدم المذكرات في قاموس لونغمان المعاصر للإنجليزية أو قاموس أكسفورد المتقدم للمتعلمين) حول كيفية استخدامها صحيحة.
- **production, produce, product**
- **continual, continuous**
- **comprise, compose, consist of, constitute, include**
- Future time expression over the years.
- الوقت المستقبلي تعبير على مدى السنوات
- A reported feature of the recent history of English has been the rise of the use of BE going to... to express future time.
- ثمة ميزة ذكرت من التاريخ الحديث للإنجليزية صعود استخدام BE.. زمن صريح للمستقبل
- You could look at some texts of current English, older writers (e.g. Dickens) and even Shakespeare to see how often this expression occurs, and how often it seems to be used in a future time sense (not just literally to mean 'move towards').
- يمكن أن نلقي نظرة على بعض نصوص الإنجليزية الحالية، الكاتب الأكبر سنا (مثل ديكنز)، وحتى شكسبير لنرى كيف غالبا ما يحدث هذا التعبير، وكيف غالبا ما يبدو لاستخدامها في وقت شعور لاحق (وليس حرفيا فقط على أنها تعني 'نحو التحرك').
- What characters use it? (Innovations often start socially 'from below').
- ما هي أحرف استخدامها؟ (غالبا ما تبدأ الابتكارات اجتماعيا 'من أسفل').

- And perhaps you could look at one or more other means of conveying future time in the same texts.
- وربما كنت يمكن أن ننظر في واحد أو أكثر من وسيلة أخرى لنقل الزمن في المستقبل في النصوص نفسها.
- Note you would need to try and match up styles of text as far as possible across the years.
- ملاحظة قد تحتاج إلى محاولة وتتطابق من أنماط النص قدر الإمكان عبر السنين.
- Refer also to Quirk et al. or Leech on time and tenses in English.
- تشير أيضا إلى المراوغة وآخرون. أو المستنزف في وقت محدد وأزمة في الإنجليزية
- Frequency and a 'lexical syllabus' for learners
- التردد و'المنهج المعجمي' للمتعلمين
- Syllabus makers have often attempted to control the introduction of vocab items in a language course, and the most popular criterion has been frequency. I.e. the course introduces new vocab roughly in order of decreasing frequency in the target language, based on some count.
- وكثيرا ما حاول صناع المنهج إلى التحكم في مقدمة بنود فوكب في دورة اللغة، وكان المعيار الأكثر شعبية التردد. أي يقدم فوكب هذا المساق الجديد تقريبا لأجل تقليل تردد في هدف اللغة ، استنادا إلى بعض العد.
- The most popular count relied on in EFL from the 40s onwards was Michael West's General Service List of English Words.
- كان العد الأكثر شعبية الاعتماد عليها في EFL ابتداء من S40 قائمة مايكل الغرب من فئة الخدمات العامة من الكلمات الإنجليزية.

- Recently this idea has received a new lease of life under the banner of the 'lexical syllabus', and today we have counts based on far larger corpora than West's 5 million or so.
- مؤخرا تلقت هذه الفكرة فرصة جديدة للحياة تحت راية "المنهج المعجمي"، واليوم لدينا حساب على أساس corpora أكبر بكثير مما في الغرب 5 ملايين أو نحو ذلك.
- See books by Willis and Lewis, and the COBUILD English Course.
- راجع كتب من قبل ويليس ولويس، ودورة الانجليزية
- You could take a course book which lists the new words in it, or a syllabus which provides lists of words to be known by different levels, and see how far they seem to be selecting and grading in accord with frequency.
- كنت قد تأخذ الكتاب بالطبع الذي يسرد كلمات جديدة بذلك، أو المنهج الذي يقدم قوائم من الكلمات لتكون معروفة من قبل مستويات مختلفة، ونرى إلى أي مدى يبدو أنها اختيار وتدرج في اتفاق مع التردد.
- You would have to sample the items and check their frequency and produce a profile.
- سيكون لديك لأخذ عينات من المواد والتحقق من وتيرتها وانتاج الصفحة الشخصية.
- Or scan them and use the Complete Lexical Tutor online.
- أو مسحها ضوئيا واستخدام كامل المعجم كمعلم على الانترنت.
- If not selected by frequency, by what criteria then?
- إذا لم يتم التحديد حسب التردد، ما هي المعايير بعد ذلك؟