



# LANGUAGE AND IT

## اللغة وتقنية المعلومات

DR. ABDULAZIZ ALTURKI

د. عبد العزيز التركي

إعداد: شرقاوي

تنسيق : FEDWA HASAN

### How did IT change language?

#### كيف غيرت تكنولوجيا المعلومات اللغة؟

- IT: (information technology)
  - CALL : (IT IS Learning Language assisted computer )
  - Technology and human communication (past vs. present)
  - How do technologists view the future of human-machine communication? (Artificial intelligence?)
  - Technology and education ? (Blackboard)
  - Distance learning?
  - (technologists – social pragmatists – English teachers' views about the use of IT).
- 
- ( IT ) تكنولوجيا المعلومات
  - ( CALL ) تعلّم اللغة بمساعدة الكمبيوتر
  - التكنولوجيا والاتصال البشري (الماضي مقابل الحاضر)
  - كيف ينظر الفنيون إلى مستقبل التواصل بين الإنسان والآلة؟ (الذكاء الاصطناعي؟)
  - التكنولوجيا والتعليم. (بلاك بورد)
  - الدراسة عن بعد.
  - (التقنيين – البراغماتيين- الاجتماعيين) آراء مدرسي اللغة الإنجليزية حول استخدام تكنولوجيا المعلومات.
- 
- **Pragmatists:** The people who think of society and negative aspects or changes in human communication. And they concern about the future communication.
  - **البراغماتيون:** الناس الذين يفكرون حول المجتمعات والجوانب السلبية في التواصل الإنساني، ويهتمون بالاتصالات في المستقبل.
  - **The aim of the ELT in the past:** To improve the student's communication competence by sending the student to a Country that have native language.(emigration)
  - **هدف تعليم اللغة الانجليزية في الماضي:** لتحسين كفاءة الطالب التواصلية، عن طريق ارسال الطالب إلى البلد ذات اللغة الأصلية (هجرة).

- Use of IT in language learning
- Practice ! (past vs. present)
- Aim of ELT (students' communicative competence?)
- Communicative competence in face-to-face vs. IT communication (rules)

- استخدام تكنولوجيا المعلومات في تعلم اللغة
- الممارسة ! (الماضي مقابل الحاضر)
- الكفاءة التواصلية للطلاب؟ ELT (الهدف من)
- الكفاءة التواصلية في التواصل وجهاً لوجه مقابل تكنولوجيا المعلومات (قواعد)

## How did IT change EL?

### 1. English is changing. commutative way

- Language contact (loan words – music – food..etc).
- Netspeak : Oxford dictionary added 65 new words this month (FOMO – emoji – internet of things –phablet – selfie ) / LOOL-Laughing outlandishly out loud-BRB-google something...etc)
- New inventions: Ipad – fax - email
- No gatekeeping.
- Language became informal and it is on line

## كيف تغيرت تكنولوجيا المعلومات؟

### 1. الإنجليزية نفسها تتغير

- بطريقة التواصل بواسطة لغة المحتوى
- وفي الكلمات المستعارة-الموسيقى-الغذاء مثلاً
- وفي كلام الانترنت : قاموس أكسفورد وأضاف 65 كلمة جديدة هذا الشهر مثل ( FOMO – emoji – internet of things –phablet – selfie ) / LOOL-Laughing outlandishly out loud-BRB-google something...etc)
- الاختراعات الجديدة : آيباد-فاكس-البريد الالكتروني
- الرقابة أقل.

ونتيجة لذلك أصبحت اللغة الانجليزية غير رسميه ومقبولة.

## 2. IT and the study of language.

- The study of phonology – speech recognition systems
- Grammar – Authoritarian vs. corpus based grammar.
- Corpus linguistics.

## 2. تكنولوجيا المعلومات ودراسة اللغة

- دراسة علم الصوتيات
- نظام التعرف علي الكلام
- تأليف القواعد مقابل مجموعة القواعد الأساسية

## 3. Tasks for ELT.

Traditional vs. contemporary tasks. (will be explored in detail in other lectures).

CMC + LC interaction

a. Computer-mediated-communication:

Synchronous vs. asynchronous communication software in classes.

a. Learner-computer interaction:

CMC vs. learner-computer interaction

L-C interaction (hyperlinks – word definition balloons- corrective feedback..etc)

## 3. مهام من أجل تعلم اللغة:

المهام التقليدية مقابل المعاصرة. (سيتم استكشافها بالتفصيل في محاضرات أخرى)

الاتصال بواسطة الكمبيوتر+تفاعل الكمبيوتر مع الطالب.

a. الاتصال بواسطة الكمبيوتر:

متزامن في مقابل برامج الاتصالات غير المتزامنة في الفصول الدراسية

b. تفاعل الكمبيوتر مع الطالب:

التفاعلات بين الطالب والكمبيوتر مثل (ارتباطات تشعبية - تعريف بالونات الكلمة - ملاحظات تصحيحية .. الخ)

#### 4. New forms of assessments.

IT-based assessments tasks (double-edged sword)?

Listening exams validity (no video) vs. computer exams

Note-taking in a computer-delivered test (easy?)

Highlighting a reading text portion in CBT TOEFL.

Economics !

#### 1. أشكال جديدة للتقييمات:

مهام تقييم تستند إلى تكنولوجيا المعلومات (سلاح ذو حدين).  
صلاحية اختبارات الفهم الاستماعي مقابل اختبارات الكمبيوتر.  
تدوين الملاحظات في اختبار يتم تسليمه بواسطة الكمبيوتر (سهل)  
تسليط الضوء على جزء نص القراءة .  
اقتصاديات.

#### Computer Mediated Communication(CMC) :

it is a communication that mediated by computer.

(L C) : interaction computer with learner.

#### Interaction:

1. **Synchronous** :means that the communication is taking place in real time.

for example, sit in the computer lab during the course period to read and respond to each other's messages discussing a story that they have read.

2. **Asynchronous**: communication allows learners to read/speak and write/hear electronic messages, which are stored on a server to be produced and accessed any time.

#### (CMC الاتصالات بواسطة الكمبيوتر)

هي عبارة عن اتصال عن طريق الكمبيوتر.

(L C) : تفاعل الكمبيوتر مع المتعلم

هذا التفاعل نوعين :

1- **متزامن**: يعني أن الاتصالات تجري في الوقت الفعلي (مباشر) ،

على سبيل المثال ، جلوسهم في مختبر الكمبيوتر خلال فترة الدورة التدريبية للقراءة والاستجابة لرسائل بعضهم البعض ومناقشة قصة قرأوها.

2- **غير متزامن**: يتيح التواصل للمتعلمين قراءة / التحدث وكتابة / سماع الرسائل الإلكترونية ، التي يتم تخزينها على خادم يتم إنتاجها والوصول إليها في أي وقت

## 5. Research on learning:

- (Learner language) Data from computer mediated tasks can be studied by researchers to examine strategies, errors, style..etc.
- Teacher education and applied linguistics:

Students need to develop an understanding of fundamental issues and concepts in applied linguistics (communicative competence – learner language- registers- communication tasks)

- IS CALL a departure point from applied linguistics, or a continuation of the field?  
No
- Technology without applied linguistics ?
- Teachers need to learn how to use IT to construct and implement materials for teaching & assessing English.

### 5. بحوث حول التعلّم:

- (لغة المتعلم) يمكن دراسة البيانات المأخوذة من مهام الكمبيوتر بوساطة الباحثين لدراسة الاستراتيجيات والأخطاء والأسلوب...إلخ
  - تعليم المعلمين واللغويات التطبيقية:
- يحتاج الطلاب إلى تطوير فهم القضايا والمفاهيم الأساسية في اللغويات التطبيقية (الكفاءة الاتصالية - اللغة المتعلمة - السجلات - مهام الاتصال)
- هل التكنولوجيا نقطة منفصلة عن علم اللغويات التطبيقية ، أم استمرار ؟ لا ..انما هي استمرارية له.
  - يحتاج المدرسون إلى تعلم كيفية استخدام تكنولوجيا المعلومات لإنشاء وتنفيذ مواد لتعليم اللغة الإنجليزية وتقييمها.

## Lecture 2

### The potential of IT for language learning

#### امكانيات تكنولوجيا المعلومات لتعلم اللغة

#### The potential of technology for language learning.

- Number of ELT webpages.
- Effective?

#### إمكانيات التكنولوجيا لتعلم اللغة.

- عدد صفحات الويب. ELT.
- فعال؟

#### Language learning and instruction. p35.

- Traditional vs. virtual immersion
- 'Comprehensible input' - Krashen (1982) :
- (Language comprehended without the learner knowing all of the linguistic forms in the message).

Q- Is CI sufficient to explain learning & linguistic development?

NO ..... spoken language/ grammar? p.36

How can e-learning help.

#### تعليم اللغة وتعليمها. P35.

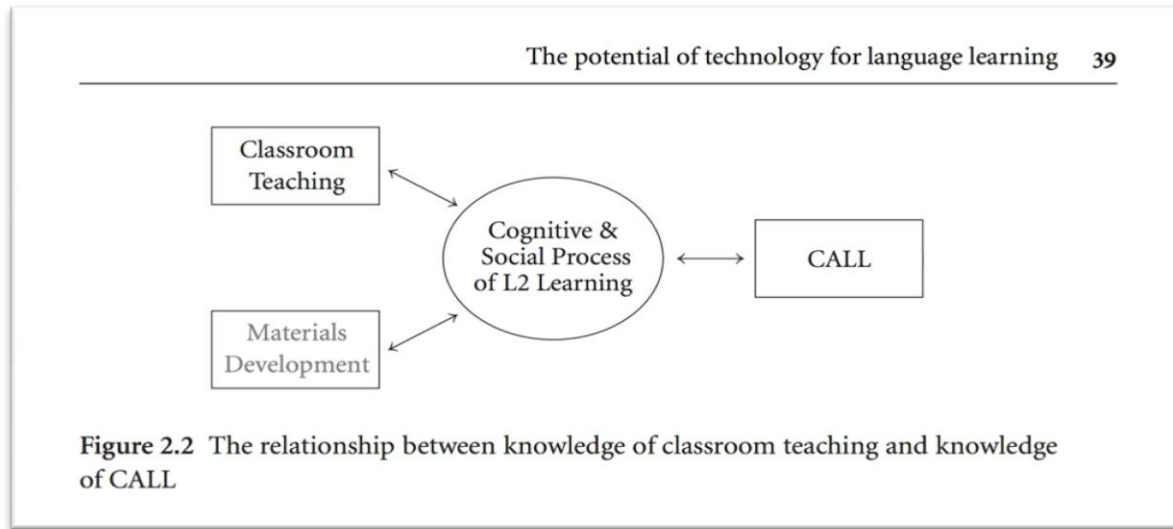
- الغمر التقليدي مقابل الظاهري
- "الإدخال المفهوم: (1982) Krashen - "
- (يفهم اللغة دون المتعلم مع معرفة جميع الأشكال اللغوية في الرسالة).

هل تكفي CI لشرح التعلم والتطوير اللغوي؟

P.36 لا ... اللغة المنطوقة / القواعد؟

كيف يمكن أن يساعد التعلم الإلكتروني

## Insights from theory and research:



...with respect to their mutual interests in the cognitive and social processes of L2 learning. . .  
From the cognitive perspective, among the most prominent [interests] are L2 comprehension, planning, and production; motivation; and attention to, and awareness of, L2 meaning and form. Social processes include various forms of communication and interaction, ranging from collaborative dialogue to instructional intervention, with mediation through negotiation of meaning. (Pica 1997: 56)

Learning = social (classroom) activities and tasks + cognitive processes.

### رؤى من النظرية والبحث:

. فيما يتعلق بمصالحهم المشتركة في العمليات المعرفية والاجتماعية للتعلم . . . L2. من المنظور المعرفي ، من بين أبرز [الاهتمامات] الفهم والتخطيط والإنتاج L2 ؛ التحفيز؛ والانتباه إلى ، والوعي ، والمعنى L2 والشكل. تشمل العمليات الاجتماعية أشكالاً مختلفة من التواصل والتفاعل ، تتراوح من الحوار التعاوني إلى التدخل التعليمي ، مع الوساطة من خلال التفاوض على المعنى. (بيكا 1997: 56)

التعلم = الأنشطة الاجتماعية (الصفية) والمهام + العمليات المعرفية.



## Social and cognitive research.

### Tasks and cognitive research - 38-39.

- The need of learners to comprehend linguistic input
- The need to notice gaps between their linguistic knowledge and the target language.
- motivation is important for making the cognitive effort to engage in tasks.
- Gap noticing is prompted by requiring learners to produce the target language, and it is enhanced when learners have time to plan their production.

البحوث الاجتماعية والمعرفية: 39

### المهام والبحوث المعرفية - 38-

- حاجة المتعلمين لفهم المدخلات اللغوية
- الحاجة إلى ملاحظة الفجوات بين معرفتهم اللغوية واللغة المستهدفة
- الدافع مهم لجعل الجهود المعرفية للمشاركة في المهام
- تتم مطالبة الفجوة من خلال مطالبة المتعلمين بإنتاج اللغة المستهدفة ، ويتم تحسينها عندما يكون لدى المتعلمين وقت لتخطيط إنتاجهم.

### Tasks and social research. p39.

Collaboration between learners is seen as a key to development. (scaffolding by interlocutors)

The context is important in constructing the identity of the learner (participant vs. marginal person).

### المهام والبحث الاجتماعي. P39 .

ينظر إلى التعاون بين المتعلمين على أنه مفتاح التنمية. (سقالة بواسطة محاورين)

السياق مهم في بناء هوية المتعلم (مشارك مقابل شخص هامشي)

## CALL tasks and cognitive/social perspectives.

### 1- The cognitive perspective

#### Enhanced input. p40.

Input = learning ?

Q- What is the role of input salience?

The likelihood of learners' acquiring linguistic input increases if their attention is drawn to salient linguistic features (40).

How ? 'input enhancement'

#### المهام CALL وجهات النظر الإدراكية / الاجتماعية.

1. المنظور المعرفي

تعزيز المدخلات. P40.

المدخلات = التعلم؟

ما هو دور المدخلات البارزة؟

تزداد احتمالية حصول المتعلمين على مدخلات لغوية إذا ما تم توجيه انتباههم إلى سمات لغوية بارزة (40).

كيف ؟ "تحسين الإدخال"

Table 2.1 Types of enhanced input expected to be beneficial to learners

Input Enhancement	Description
Salience	Marking a grammatical form on the screen or phonologically through stress Repeating a grammatical form or lexical phrase
Modification	Making the input understandable to the learner through any means that gets at the meaning (e.g., images, L1 translation, L2 dictionary definitions, simplification)
Elaboration	Increasing the potential for understanding the input through addition of plausible, grammatical L2 elaborations to the original text (e.g., defining relative clauses)

Input without any forms marked:

Rabbits like to eat many different types of garden flowers, particularly when the plants are young. People attempt to save their plants by placing substances in the garden that the rabbits do not like. Moth balls, human hair, and marigolds may help to deter hungry rabbits. However, the only certain method is a good fence.

Input with to-complements made salient:

Rabbits like to eat many different types of garden flowers, particularly when the plants are young. People attempt to save their plants by placing substances in the garden that the rabbits do not like. Moth balls, human hair, and marigolds may help to deter hungry rabbits. However, the only certain method is a good fence.

Figure 2.3 Texts with (on the right) and without (on the left) highlighted forms

Click to hear the underlined words.

Rabbits like to eat many different types of garden flowers, particularly when the plants are young. People attempt to save their plants by placing substances in the garden that the rabbits do not like. Moth balls, human hair, and marigolds may help to deter hungry rabbits. However, the only certain method is a good fence.

Choose the best answer based on the text.

Rabbits like

- ☐ young plants
- ☐ substances
- ☐ marigolds

The best way to deter rabbits is with

- ☐ moth balls
- ☐ marigolds
- ☐ fences

Figure 2.4 A task prompting vocabulary repetition

## 2. Input modification. p45.

What is it?

modifications appear as hypertexts or hypermedia links that help the learners to comprehend the input.

### 3. تعديل المدخلات. P45.

ما هذا؟

تظهر التعديلات على شكل نصوص تشعبية أو روابط hypermedia تساعد المتعلمين على فهم المدخلات.

- **Types of modification**

- a. Images

- أنواع التعديل

- a. الصور

48 Chapter 2


<p>On the weekend, Martha liked to stop by to see her father at work where she was greeted by the <u>Dalmatian</u> that lived at the station. Lunch was served at 12:30 every day so she tried to get there in time to eat if she could. But last Saturday, when lunch was served, Martha was not there. Her ...</p>	 <p>On the weekend, Martha liked to stop by to see her father at work where she was greeted by the <u>Dalmatian</u> that lived at the station. Lunch was served at 12:30 every day so she tried to get there in time to eat if she could. But last Saturday, when lunch was served, Martha was not there. Her ...</p>
--	--

Figure 2.7 Before (left) and after (right) the learner clicks on a hypermedia link containing an image depicting the word “Dalmatian”

## Types of modification

- b. L1 translation.

- b. الترجمة.

The potential of technology for language learning 49

**To tan or not to tan**

Two researchers specializing in the psychology of health say they’ve found a more productive way to wean sun worshipers from catching más pálido. They’ve proven that when you actually show people what ultraviolet (UVA) is doing to skin, they have a surprisingly high tendency to settle for the pastier look.

(by Bridget Bailey, *Inside Iowa State*, August 30, 2002)

Figure 2.8 Modifications through L1 translation in Spanish

c. L2 definitions.

c. تعريفات اللغة الثانية.

d. simplification.

d. بالتبسيط.

50 Chapter 2

Original text:

To tan or not to tan

Two researchers specializing in the psychology of health say they've found a more productive way to wean sun worshipers from catching some rays. They've proven that when you actually show people what ultraviolet (UV) radiation is doing to skin, they have a surprisingly high tendency to settle for the pastier look.

Simplifications:

Should people get a tan?

Two psychology professors said that they made an important discovery. They found out how to keep people away from the sun. The professors show the people pictures of the effects of the sun on their skin. Then these people choose to stay out of the sun.

(by Bridget Bailey, *Inside Iowa State*, August 30, 2002)

Figure 2.9 An authentic text (left) and a simplified version (right)

Input elaboration.

Original text:

To tan or not to tan

Two researchers specializing in the psychology of health say they've found a more productive way to wean sun worshipers from catching some rays. They've proven that when you actually show people what ultraviolet (UV) radiation is doing to skin, they have a surprisingly high tendency to settle for the pastier look.

Elaborated text:

To get a tan or not to get a tan

Two researchers who specialize in the psychology of health say they've found a more productive way to wean, or prevent, sun worshipers from catching some rays, keeping them away from the sun. The researchers have proven that when they actually show people what ultraviolet (UV) radiation is doing to skin, the sun lovers have a surprisingly high tendency to settle for the pastier look rather than getting a tan.

(by Bridget Bailey, *Inside Iowa State*, August 30, 2002)

Figure 2.10 Elaborated input (right) developed from an authentic text (left)

- **Input modification** : refers to the provision of an accessible rendition of the L2 input.
- **Modifications**: can be any form of simplification, repetition, clarification, or L1 translation – anything that an interlocutor does during the course of a conversation to clarify meaning in order to continue a conversation.
- **Types of modification**:
  - images
  - translation L2
  - definitions L2
  - Simplification

- **تعديل المدخلات**: يشير إلى توفير ترجمة يمكن الوصول إليها من مدخلات اللغة الثانية.
- **التعديلات**: يمكن أن يكون أي شكل من أشكال التبسيط ، التكرار ، التوضيح ، أو ترجمة الى اللغة الأولى أو أي شيء يفعله المحاور أثناء المحادثة لتوضيح المعنى من أجل مواصلة محادثة.
- **أنواع التعديلات**:

- بالصور
- بالترجمة باللغة الأولى
- بتعريفات باللغة الثانية
- بالتبسيط

**Input elaboration** : Input elaboration is intended to help learners gain access to the meaning of the text by adding grammatical phrases and clauses such as defining appositives, relative clauses, and restatements.

**تفصيل المدخلات**: يهدف تصميم المدخلات إلى مساعدة المتعلمين على الوصول إلى معنى النص عن طريق إضافة العبارات النحوية والبنود مثل تحديد الإيجابيات ، والبنود النسبية ، وإعادة التصحيحات.



### Evaluating course materials

#### تقييم المواد الدراسية

#### Interaction in CALL.

Interpersonal communication:

#### The benefits of interaction among learners;

- 1- negotiation of meaning
- 2- co-constructing meaning and
- 3- prompting learners' attention to form.

#### التفاعل في (تعلم اللغة بمساعدة الكمبيوتر)

التواصل بين الأشخاص

#### فوائد التفاعل بين المتعلمين:

1. التفاوض على المعنى
2. المشاركة في بناء المعنى
3. حث المتدربين على الاهتمام بالشكل.

#### Example from CMC: في الجدول أمثلة على الاتصال بواسطة الكمبيوتر.

Table 2.3 Negotiation of meaning in synchronous communication

Participant	Language
Instructor	What do you think about sustaining life artificially?
Student 1	What is "sustaining artificially"? Anyone answer me.
Student 2	what's that???
Student 3	Artificially support someone's life!
Student 4	Don't ou understand artifically?
Student 3	For example using machines!

(from Rodriguez 1998)

First move: Computer shows text	Second move: Learner clicks on word	Third move: Computer shows definition
BRUSSELS, Belgium – They had to struggle a bit to open some recalcitrant bottles of champagne, but finance ministers from 11 countries celebrated New Year's Eve Thursday by launching the euro as the continent's new unified currency.	BRUSSELS, Belgium – They had to struggle a bit to open some <b>recalcitrant</b> bottles of champagne, but finance ministers from 11 countries celebrated New Year's Eve Thursday by launching the euro as the continent's new unified currency.	recalcitrant – Definition: formal adjective; refusing to obey or be controlled, even after being punished: recalcitrant behavior BRUSSELS, Belgium – They had to struggle a bit to open some <b>recalcitrant</b> bottles of champagne, but finance ministers from 11 countries celebrated New Year's Eve Thursday by launching the euro as the continent's new unified currency.

Figure 2.12 Three moves in an interaction requesting a word definition (see Hegelheimer & Chapelle 2000 for an interactive on-line version)

### Production in CALL tasks.

### الإنتاج في مهام تعلم اللغة بواسطة الكمبيوتر

Production and planning/ correction.

الإنتاج والتخطيط / التصحيح

### Good vs. bad CALL materials.

### الجيد مقابل الرديء في مواد تعلم اللغة بواسطة الكمبيوتر.

Now you can evaluate CALL materials.

الآن يمكنك تقييم مواد تعلم اللغة بواسطة الكمبيوتر.

### Evaluating language learning:

- Again, the number of CALL materials and their effectiveness.
- Intuition & courseware development
- Solution? careful evaluation of instructional effectiveness -(research)- MacWhinney.

### تقييم تعلم اللغة:

- -مرة أخرى عدد مواد تعلم اللغة بواسطة الكمبيوتر وفعاليتها.
- -الحدس وتطوير البرامج التعليمية.
- حل/ تقييم دقيق للفعالية التعليمية - (البحث).



**But who is interested about the effectiveness of using a particular software in teaching?**

- L2 teachers
- Administrators
- Publishers
- (Role play) -What are their views on CALL & its effectiveness?

ولكن من هم الأكثر اهتماما ببرمجيات معينه في التعليم؟

- مدرسين
- المسؤولين
- الناشرين
- (تمثيل الأدوار) - ما هي وجهات نظرهم بشأن تعلم اللغة بواسطة الكمبيوتر وفعاليتها؟

## 76 Chapter 3

**Table 3.1** The contradictions in discourse about CALL vs. classroom research

A case must be made for CALL	My experience
Language teachers need to be convinced about whether to use CALL and for what purpose.	The results of research comparing “the computer” to “the classroom” are not conducive to developing principles of language learning and teaching.
Administrators need CALL vs. classroom comparisons to be convinced to fund computer labs.	Administrators seem to make decisions about funding computer labs on the basis of societal or institutional factors rather than of comparison research.
Publishers need positive results from CALL vs. classroom research to sell their software.	Current professional discourse teaches teachers that they need the software that publishers produce.

- **Aspects of production theory:**

First, from a cognitive perspective, the benefits of producing language may be enhanced when learners have the opportunity to plan before speaking or writing.

Second, the cognitive view also emphasizes the importance of opportunities to correct linguistic output, which can be prompted by feedback from others or from self-evaluation

Third, the sociocultural perspective suggests the value of the learners' use of the help of the interlocutor to allow for production beyond what the learner could accomplish alone.

- **جوانب نظرية الإنتاج:**

أولاً ، من منظور إدراكي ، يمكن تحسين فوائد إنتاج اللغة عندما تتاح للمتعلمين الفرصة للتخطيط قبل التحدث أو الكتابة.

ثانياً ، تؤكد النظرة المعرفية أيضاً على أهمية الفرص لتصحيح المخرجات اللغوية ، والتي يمكن أن تدفعها ردود فعل من الآخرين أو من التقييم الذاتي.

ثالثاً ، يشير المنظور الاجتماعي والثقافي إلى قيمة استخدام المتعلمين لمساعدة المحاور للسماح بالإنتاج بما يتجاوز ما يمكن للمتعلم تحقيقه بمفرده.

### **What is the best way to check materials effectiveness?**

- RESEARCH !
- But which methodology ? Qualitative vs. quantitative ?
- The role of research questions
- Chapelle (1997): The following must be addressed
  1. What kind of language does the learner engage in during CALL activity?
  2. How good is the language experience in CALL for L2 learning?

### **ما هي أفضل طريقة للتحقق من فعالية المواد؟**

- الأبحاث
- ولكن أي منهجية؟ النوعية مقابل الكمية؟
- دور أسئلة البحث
- Chapelle (1997): يجب تناول ما يلي
  1. ما نوع اللغة التي يشارك بها المتعلم أثناء نشاط CALL
  2. ما مدى جودة تجربة اللغة في CALL for L2 learning

## Examples of useful CALL research:

### What is useful CALL research? p.81

Research that provides some evidence about the design of the software, the learners' use of CALL, or the way that the teacher has organised the task.

### أمثلة على أبحاث CALL المفيدة

#### ماهي فائدة بحث CALL ؟

البحوث التي تقدم بعض الأدلة حول تصميم البرنامج ، واستخدام CALL للمتعلمين ، أو الطريقة التي نظم بها المعلم هذه المهمة.

82 Chapter 3

Table 3.2 Three approaches to developing useful research questions about CALL

Focus on. . .	Results	Primarily for
The software	Indicate the most successful software design strategies	Software developers and lab coordinators
The learner	Indicate successful strategies for using software	Teachers and students
The task	Indicate the best ways to structure learning tasks	Teachers

### 1. Focus on software:

#### A. Subtitles for listening

- Borrás and Lafayette (1994) investigated the effectiveness of optional subtitles as a means of modifying the input. L2 French.
- Two groups (optional subtitles vs. no subtitles).
- Speaking task followed.
- The results clearly favoured the subtitle group.

#### 1. التركيز على البرامج:

##### أ- ترجمات للاستماع

- حقق بوراس و لافاييت (1994) في فعالية الترجمات الاختيارية كوسيلة لتعديل المدخلات L2 الفرنسية.
- مجموعتين (ترجمات اختيارية مقابل عدم وجود ترجمات)
- تحدثت مهمة التحدث.
- فضلت النتائج بوضوح مجموعة الترجمة

## B. Intelligent feedback for grammar (L2 Japanese)

- Intelligent vs. unintelligent feedback groups (2)
- The intelligent condition was favoured in both post-tests and end of semester tests.

### B. ملاحظات ذكية لقواعد اللغة (L2 Japanese)

- مجموعات ردود الفعل الذكية مقابل غير ذكية (2)
- تم تفضيل الحالة الذكية في كل من الاختبارات اللاحقة واختبارات نهاية الفصل الدراسي

## 2. Focus on the learners:

- a. Looking up words (modified input)
- b. Asking for help
- c. Participating in telecollaboration

### 2. التركيز على المتعلمين:

- أ- البحث عن الكلمات (المدخلات المعدلة)
- ب- طلب المساعدة
- ت- المشاركة في مجال التعاون عن بعد

## 3. Focus on the learning task.:

A- Web-based listening (incidental vocabulary acquisition (Kon 2002). Pre-test & Post-test.

### 3. التركيز على مهمة التعلم

A- الاستماع عبر الإنترنت (اكتساب المفردات العرضية. (Kon 2002) اختبار ما قبل وبعد الاختبار

Table 3.4 Analysis of input modes and success of acquisition for vocabulary (from Kon 2002: 52)

Mode of presentation	Number of modes	% of words acquired*	Quality of the input for acquisition
Audio-video	1	25	OK
Audio-video & written comprehension questions	2	32	Better
Audio-video & overhead transparency notes	2	39	Better
Audio-video, written comprehension questions, and overhead transparency notes	3	67	The best

\* Based on delayed posttest performance

Web-based listening activity can facilitate incidental vocabulary acquisition. - Modes & repetition also help

- B- Text chat as rehearsal. p.89
- Compton (2002) investigated a text chat-based task as a means of increasing students' willingness to communicate through oral language in the classroom.
- Some learners benefit from the opportunity to engage in the texts chat before engaging in oral classroom work. (planning & production).

• ب- الدردشة النصية والبروفة P.89 .

- حقق كومبتون (2002) في مهمة قائمة على المحادثة النصية كوسيلة لزيادة رغبة الطلاب في التواصل من خلال اللغة الشفوية في الفصل الدراسي.
- يستفيد بعض المتعلمين من الفرصة للمشاركة في نقاش الدردشة قبل التعامل مع الفصل الدراسي الشفوي. (التخطيط والإنتاج)

**The role of theory.**

دور النظرية

### Investigating learners' use of IT

#### تحقيق استخدام المتعلمين لتكنولوجيا المعلومات

#### التحقيق في استخدام المتعلمين للتكنولوجيا. Investigating learners' use of technology

- Previous chapters (ELT is changing - the potential of technology for ELT- evaluating language learning)
- One of the two useful RQs : 'What kind of language does the learner engage in during CALL activity?'

Investigating learners' use of technology (process data) p.97

(traditional research vs. CALL process data research) and attention

- الفصول السابقة ELT ( آخذة في التغير - إمكانات التكنولوجيا لتقييم - تعلم اللغة )
  - واحد من اثنين من المنافع المفيدة: "ما نوع اللغة التي يشارك فيها المتعلم أثناء نشاط CALL ؟
- استقصاء استخدام المتعلمين للتكنولوجيا (بيانات العمليات
- ( البحث التقليدي مقابل أبحاث بيانات عملية CALL ) والاهتمام

#### What is process data ? P.98

#### ما هي بيانات العملية؟

Process data is the observable record of learners' work on CALL tasks.

What are some examples of process data?

- Production of an error and receipt of intelligent feedback, a request for and receipt of translation, a linguistic production and self-correction. (computer-human interaction or CMC)
- بيانات العملية هي السجل الملحوظ لعمل الطلاب في مهام CALL .
- ما بعض الأمثلة على بيانات العمليات؟
- إنتاج خطأ واستلام ردود فعل ذكية ، وطلب واستلام الترجمة ، وإنتاج لغوي وتصحيح ذاتي. ( التفاعل بين الكمبيوتر والبشرية أو CMC )

Text 2. Learners engaging in a chat through written language

Alloy (07:41:08): Hi, Arosio, have you been to Malaysia & Borneo Island.

Naghib (07:41:18): What about YOUR men Carmilla?

Alloy (07:42:01): I am from Kuohing city, in Sarawak in Borneo Island. It's part of Malaysia.

Richard M. (joined the chat at 7:42:08)

Arosio (07:42:26): Speed: wake up!!!

Naghib (07:42:36): Sppe is right: it is getting very boring here.  
Let's speak of sth more interesting. . .

(Negretti 1999: Appendix A)

Text 3. A Text from a listening comprehension task

Computer: I'm looking for a job as a waiter

Learner: <clicks on "waiter">

Learner: <clicks on the vocabulary button>

Computer: waiter

DEFINITION

a person who serves customers at tables in a restaurant.

EXAMPLE

Ned worked as a waiter before going into acting. (Park 1994:200)

**Ethical issue:**

Participants' permission must be gained before using process data or "moves".

**What is a move ? and who can perform it? p.101**

Each unit of behaviour, called a move, can consist of either language or behaviour, and can be performed by either the learners or the computer

**قضية أخلاقية:**

يجب الحصول على إذن المشاركين قبل استخدام بيانات العملية أو "التحركات".

**ما هي الخطوة؟ ومن يستطيع القيام بها؟ P.101**

يمكن أن تتكون كل وحدة من وحدات السلوك ، تسمى حركة ، إما من لغة أو سلوك ، ويمكن تنفيذها من قبل المتعلمين أو الكمبيوتر.



## Text 2

Alloy	Naghib	Alloy	Richard M	Arosio	Naghib
07:41:08	07:41:18	07:42:01	7:42:08	07:42:26	07:42:36
Hi, Arosio, have you been to Malaysia & Borneo Island.	What about YOUR men Carmilla?	I am from Kuohing city, in Sarawak in Borneo Island. It's part of Malaysia.		Speed: wake up!!!	Sppe is right: it is getting very boring here. Let's speak of sth more interesting. ...

## Text 3

Computer	Learner	Learner	Computer
I'm looking for a job as a waiter <printed on screen>	<clicks on "waiter">	<clicks on the vocabulary button>	Waiter; DEFINITION a person who serves customers a tables in a restaurant; EXAMPLE Ned worked as a waiter before going into acting.

**Is that enough?**

There is a need for a theoretical perspective to explain such interactions.

**Description of PD.**

What are some methodologies used in describing PD? p.102

Interaction analysis - Discourse analysis - Conversation analysis.

هل هذا كاف؟

هناك حاجة لمنظور نظري لشرح مثل هذه التفاعلات.

وصف PD.

ما هي بعض المنهجيات المستخدمة في وصف التطوير المهني؟ p.102

تحليل التفاعل - تحليل الخطاب - تحليل المحادثة



## Interaction analysis.

### What is interaction analysis?

Interaction analysis is used to document the particular moves that the learner makes while working with technology. p.103

تحليل التفاعل.

ما هو تحليل التفاعل؟

يستخدم تحليل التفاعل لتوثيق التحركات الخاصة التي يقوم بها المتعلم أثناء العمل مع التكنولوجيا. P.103.

#### 104 Chapter 4

Computer	Learner	Computer
BRUSSELS, Belgium – They had to struggle a bit to open some recalcitrant bottles of champagne, but finance	<clicks on “recalcitrant”>	recalcitrant – Definition: formal adjective; refusing to obey or be controlled, even after being punished: recalcitrant behavior

Figure 4.2 Sequence of interaction for a dictionary look up (from Hegelheimer & Chapelle 2000)

- What are some possible research questions addressed through interaction analysis?

-How frequently do learners consult the internal glossary (e.g. click-hyperlink), and how frequently do they consult the external bilingual dictionary (e.g. copy-paste or type into an online dictionary)? p.104

- ما هي بعض الأسئلة البحثية المحتملة التي تم تناولها من خلال تحليل التفاعل؟

-كم عدد المرات التي يتشاور فيها المتعلمون مع المسرد الداخلي (على سبيل المثال ، انقر فوق ارتباط تشعبي) ، وكم عدد المرات التي يتشاورون فيها مع القاموس ثنائي اللغة الخارجي (على سبيل المثال ، النسخ واللصق في قاموس على الإنترنت)؟ P.104

## Discourse analysis:

- What is DA ?

Functional description of the linguistic choices and moves that the participants make to construct a text. p.105

E.g. Learners' use of syntax (lack of concern for correctness in chat vs. oral classroom participation).

## تحليل الخطاب

- ماهو DA ؟

وصف وظيفي للخيارات اللغوية والتحركات التي يقوم بها المشاركون لبناء النص P.105 .

مثلا استخدام المتعلمين للبناء (عدم الاهتمام بالصحة في مقابل المحادثة الشفهية مقابل الفصل الدراسي)

## Conversation analysis:

- What is CA?

CA attempts to capture language users' utterances and intentions and describe how the language in discourse was used to accomplish communicative intent. p.107

- Examples?

E.g. discovering the conversational routines that the learners used to accomplish openings, closings, topic shifts, and cohesion. p.107

## تحليل المحادثة:

- ماهو CA ؟

يحاول CA التقاط الكلمات والنوايا لمستخدمي اللغة ووصف كيفية استخدام اللغة في الخطاب لتحقيق غرض . P.107

التواصل

## أمثلة:

مثلا اكتشاف الروتين التحادثي الذي استخدمه المتعلمون لإنجاز الفتحات والإغلاق والتغييرات الموضوعية . والتماسك

P.107

A process which involves interpretation of the data in a way that makes them meaningful and useful for research. P.111

عملية تتضمن تفسير البيانات بطريقة تجعلها ذات معنى ومفيدة للبحوث P.111 .

### About what ?

- -Learners' capacities
- tasks

عن ماذا ؟

- عن قدرات المشاهدين
- وعن المهام

### 1. Inferences about capacities

Inferences about learners' capacities.

E.g. if the process data were instances of dictionary checking in an electronic text, the capacities responsible might be a mental lexicon lacking the particular words that were checked. p. 114

#### 1. الاستدلالات على القدرات

استنتاجات حول قدرات المتعلمين.

مثلا إذا كانت بيانات العملية عبارة عن أمثلة من قاموس التحقق في نص إلكتروني ، قد تكون القدرات المسؤولة معجم عقلي يفتقر إلى الكلمات المحددة التي تم فحصها. ص. 114

### 2- Inferences about tasks

How the task influences learners' interaction.

E.g. using CMC (Chat) fostered the use of formal language. p. 115

#### 2. الاستدلالات حول المهام

كيف تؤثر المهمة على تفاعل المتعلمين.

(دردشة) عزز استخدام اللغة الرسمية. ص. 115 CMC مثلا استخدام

### Advancing AL: L2 learning tasks

#### L2 تطوير مهارات التعلم

##### The study of L2 learning tasks . p.128-129

- Instruction & L2 development: research findings.
- Shift of focus from TEACHING methods to tasks that LEARNERS engage in.
- Tasks must have goals and they are carried out through participants' engagement in goal-oriented behaviour that relies at least in part on language.
- Emphasis on communication tasks.

##### L2. p.128-129 دراسة مهام التعلم

- نتائج البحوث. L2 تطوير التعليم و
- تحويل التركيز من أساليب التدريس إلى المهام التي يشارك فيها المتدربون.
- يجب أن يكون للمهام أهداف وأن يتم تنفيذها من خلال مشاركة المشاركين في السلوك الموجه نحو الهدف والذي يعتمد جزئيًا على اللغة جزئيًا.
- التركيز على مهام الاتصال.

##### Advancing AL: L2 learning tasks:

The future of technology (AL).

Technology is taking over our jobs! (AL)

Knowledge of technology boosts employability.

Why? L2 tasks and work efficiency

##### L2:: تطوير مهارات التعلم من المستوى

مستقبل التكنولوجيا. (AL)

التكنولوجيا تتولى وظائفنا (AL) !

معرفة التكنولوجيا تعزز قابلية التوظيف.

لماذا؟ L2 المهام وكفاءة العمل

## How can technology help:

- 1- We will look at key elements of L2 tasks (evaluation & description).
- 2- We will consider an example of technology-mediated L2 tasks.

كيف يمكن أن تساعد التكنولوجيا:

- 1- سننظر في العناصر الرئيسية للمهام L2 (التقييم والوصف)
- 2- سننظر في مثال على مهام L2 التي تتم بوساطة التكنولوجيا

## Task evaluation. p.130:

There are three approaches to evaluate L2 tasks:

1. Outcomes
2. Negotiation of meaning
3. Three dimensions of proficiency

## تقييم المهمة p.130

هناك ثلاث طرق لتقييم مهام L2

- 1- النتائج
- 2- التفاوض على المعنى
- 3- ثلاثة أبعاد من الكفاءة

### 1. Outcomes:

- We need to assess the learning outcomes of learners who have worked on the tasks.

E.g. different vocabulary acquisition tasks. (pretest - treatment- posttest)

- A record of learners' interaction with each other is helpful.

### 1. النتائج :

- نحن بحاجة إلى تقييم نتائج التعلم من المتعلمين الذين عملوا في المهام  
مثلا مهام اكتساب المفردات المختلفة. (الاختبار القبلي - العلاج - الاختبار البعدي)-
- من المفيد تسجيل تفاعل المتعلمين مع بعضهم البعض

## 2. Negotiation of meaning:

- We need to look for instances of negotiation of meaning in the language of task participants.

- E.g. Communication breakdown - drawing attention to LINGUISTIC FORMS - resolving the problem = input acquired.

Examples.

### 2- التفاوض على المعنى:

- نحن بحاجة للبحث عن أمثلة للتفاوض على المعنى في لغة المشاركين في المهمة - على سبيل المثال انهيار الاتصال - لفت الانتباه إلى صيغ - LINGUISTIC حل المشكلة = الإدخال المكتسبة. أمثلة.

## 3- Three dimensions of proficiency:

- We need to evaluate language tasks through the criteria of accuracy, complexity and fluency.
- The goal of task-based instruction should be for learners to develop an effective balance between fluency and accuracy and to become able to increase the complexity of their linguistic production (Skehan).

### 3- ثلاثة أبعاد من البراعة:

- نحن بحاجة إلى تقييم المهام اللغوية من خلال معايير الدقة والتعقيد والطلاقة.
- يجب أن يكون الهدف من التعليم القائم على المهام هو أن يقوم المتعلمون بتطوير توازن فعال بين الطلاقة والدقة (Skehan) وأن يصبحوا قادرين على زيادة تعقيد إنتاجهم اللغوي

## L2 task description. p.131:

The objective of L2 task research is to describe tasks in such a way that teachers and researchers can choose and develop tasks that can be expected to produce the desired results.

### وصف المهمة: L2. p.131

الهدف من بحث مهمة L2

هو وصف المهام بطريقة تمكن المدرسين والباحثين من اختيار وتطوير المهام التي من المتوقع أن تؤدي إلى النتائج المرجوة.

**Table 5.3** Categories and features for a cognitively-oriented definition of task characteristics

Categories Features	Definitions
Code complexity	Syntactic and lexical complexity/load and variety
Cognitive complexity	The complexity of the topical content
Cognitive familiarity	Familiarity of topic, discourse genre, and task
Cognitive processing	Information organization, amount of “computation” required Clarity and sufficiency of given information
Communicative Stress	Degree of pressure in communication
Time pressure	How quickly the task must be done
Scale	The number of participants and length of texts
Modality	The speaking/writing vs. reading/listening contrasts
Stakes	How important it is to complete the task correctly
Control	Amount of influence participants have on the task

(Skehan 1996; Skehan &amp; Foster 2001)

**Technology mediated L2 tasks. p.136:**

Example : Focus on meaning/language in face-to-face vs. written chat tasks

**p.136: مهام توسط التكنولوجيا L2 .**

مثال: التركيز على المعنى / اللغة في مهام المحادثة الكتابية وجهًا لوجه.

## The study of L2 learning tasks

### دراسة مهام التعلم L2

#### Why TBLT?

#### Task-based language teaching

TBLT: is an approach which offers students opportunities to actively engage in communication in order to achieve a goal or complete a task.

In TBLT, students use language in tasks to improve their language skills.

#### لماذا TBLT؟

تعليم اللغة القائم على المهام:

TBLT: هو النهج الذي يوفر للطلاب فرصًا للانخراط بنشاط في التواصل من أجل تحقيق هدف أو إكمال مهمة - في TBLT يستخدم الطلاب اللغة في المهام لتحسين مهاراتهم اللغوية .

Relevant to learners' real-life language needs.

Tasks are motivating.

Tasks allow focus on form/meaning.

L2 language production is less stressful

Learner-centred

- ذات صلة بالاحتياجات اللغوية للدارسين في الحياة الواقعية

- المهام هي تحفيز

- تسمح المهام بالتركيز على الشكل / المعنى

- إنتاج لغة L2 أقل إجتهاداً.

- تركيز على المتعلم.



Before we can move forward in TBLT, a number of operational issues need to be resolved

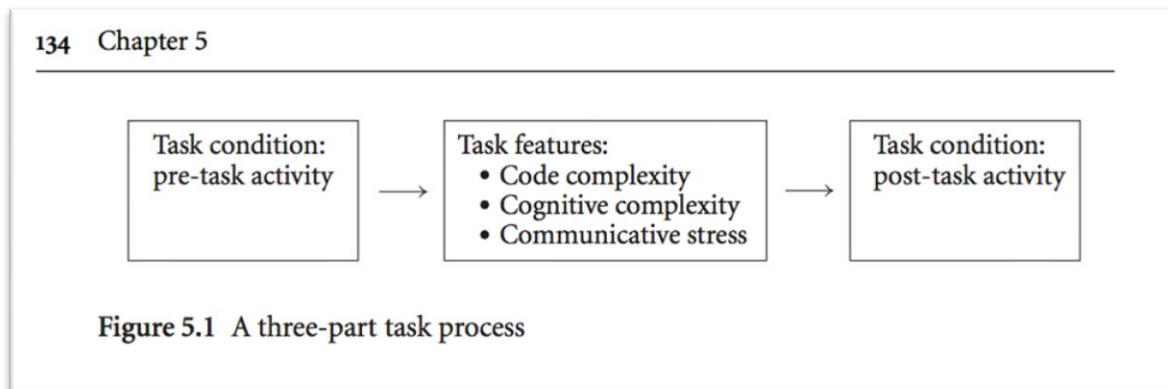
- 1- There is a need to take into account individual differences in the analysis of task-generated language.
- 2- How to empirically evaluate the overall task difficulty?
- 3- We need to take into account the conditions under which the tasks are done.

ص 133

قبل أن نتمكن من المضي قدماً في TBLT ، يجب حل عدد من المشكلات التشغيلية:

1. هناك حاجة لمراعاة الفروق الفردية في تحليل اللغة التي تم إنشاؤها في المهام
2. كيفية إجراء تقييم تجريبي لصعوبة المهمة الشاملة؟
3. نحن بحاجة إلى أن تأخذ في الاعتبار الظروف التي يتم بموجبها المهام

**Phases :** المراحل



- **pre-task phase:**

Prepares learners to perform the task linguistically and cognitively.

E.g. doing a similar task, providing a model answer, or providing planning opportunities before the task.

- **مرحلة ما قبل المهمة:**

يعد الدارسين لأداء المهمة لغوياً ومعرفياً مثلاً القيام بمهمة مماثلة ، وتقديم إجابة نموذجية ، أو توفير فرص التخطيط قبل المهمة.

- **Examples of pre-task activities:**

Students are introduced to the idea of balloon debate

Balloon debate + focus on modal verbs (If X, you might die)

Balloon debate + focus on ideas (giving each student ideas that s/he can use. (Rod Ellis)

أمثلة على أنشطة ما قبل المهمة:

يتم تعريف الطلاب على فكرة النقاش البالون:

، قد تموت (X مناقشة البالون + التركيز على الأفعال المشروطة (إذا كان

مناقشة البالون + التركيز على الأفكار (إعطاء أفكار لكل طالب يمكنه استخدامها) (رود إليس).

Advancing applied linguistics 133

**Table 5.3** Categories and features for a cognitively-oriented definition of task characteristics

Categories Features	Definitions
Code complexity	Syntactic and lexical complexity/load and variety
Cognitive complexity	The complexity of the topical content
Cognitive familiarity	Familiarity of topic, discourse genre, and task
Cognitive processing	Information organization, amount of "computation" required Clarity and sufficiency of given information
Communicative Stress	Degree of pressure in communication
Time pressure	How quickly the task must be done
Scale	The number of participants and length of texts
Modality	The speaking/writing vs. reading/listening contrasts
Stakes	How important it is to complete the task correctly
Control	Amount of influence participants have on the task

(Skehan 1996; Skehan & Foster 2001)

### **Focus on language (form) or meaning ?**

التركيز على اللغة (شكل) أو معنى؟

L2 learners find it difficult to attend to both.

This indicates a need to find a way to draw learners' attention to FORM while engaging in a communicative activity (meaning-focused).

يجد متعلمو L2 صعوبة في الحضور إلى كليهما:

يشير هذا إلى الحاجة إلى إيجاد طريقة لجذب انتباه المتعلمين إلى نموذج أثناء المشاركة في نشاط تواصل (يركز على المعنى).

**1. Implicit focus on form:**

Learner: Norah play the piano

Instructor: Norah plays the piano ?

Learner: Yes, Norah plays the piano.

**1. التركيز الضمني على الشكل:**

المتعلم: نورا تلعب البيانو

المدرّب: نورا تلعب البيانو؟

المتعلم: نعم ، تلعب نورا البيانو.

**2. Explicit focus on form:**

Learner: Yesterday, I visit my friend.

Teacher: You should use the past form 'visited'.

**2. التركيز الواضح على الشكل:**

المتعلم: بالأمس ، أقوم بزيارة صديقي.

المعلم: يجب عليك استخدام النموذج السابق "زيارة".

**3. The post-task phase:**

Task repetition (fluency, complexity, and accuracy)

Reflecting on performance (write a report)

Attention to form

**3. مرحلة ما بعد المهمة:**

تكرار المهام (الطلاقة والتعقيد والدقة)

التفكير في الأداء (كتابة تقرير)

الاهتمام بالشكل.

- **Speaking activity.**

Balloon debate - debate -

Write the rules of a country of your dreams (to foster future tense use) — Such tasks prompt students to use the future tense.. Other tasks prompt students to use different linguistic constructions.

- **How can IT help?**

**Reading/writing/listening/speaking-stylistics**

- **نشاط التحدث:**

مناقشة البالون - مناقشة-

اكتب قواعد بلد أحلامك (لتشجيع الاستخدام المتوتر في المستقبل) - هذه المهام تحث الطلاب على استخدام الزمن المستقبلي .. مهام أخرى تحث الطلاب على استخدام منشآت لغوية مختلفة.

- **كيف يمكن أن تساعد تكنولوجيا المعلومات؟**

القراءة / الكتابة / الاستماع / الناطقة الأسلوبية.

### Corpus linguistics

#### اللغويات الأساسية

- What is the most frequently used word in English ?
- What are some collocates of the word X in English?
- What adjectives collocate with the word X in English?
- The word 'tendencies' is associated with negative/positive collocates?

- ما هي الكلمة الأكثر استخدامًا في اللغة الإنجليزية؟
- في اللغة الانجليزية؟ X ما هي بعض رفق الكلمة
- في اللغة الانجليزية؟ X ما هي الصفات التي تقترن بكلمة
- ترتبط كلمة "ميل" بارتباط سلبي / إيجابي؟

#### Corpus linguistics: Definition.

#### اللغويات الأساسية : التعريف

CL is "the study of language based on examples of 'real life' language use" ..... McEnry & Wilson (2001)

1 Corpus, 2 corpora

Pre/post computer era:

هو "دراسة اللغة على أساس استخدام اللغة" CL الحياة الحقيقية (2001) McEnry & Wilson ....

1 كوربوس ، 2 كوربا

قبل / بعد عصر الكمبيوتر:

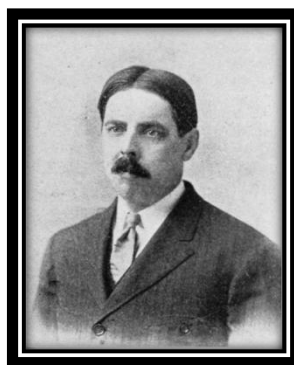
Edward Thorndike 1921

4.5 million word corpus

most frequent 30,000 words

word usage

and language change !!



إدوارد ثورندايك 1921

4.5 مليون كلمة

الأكثر شيوعا 30000 كلمة

استخدام الكلمة

وتغيير اللغة !!

- **Corpora:**

The British national corpus (BNC)

The Corpus of Contemporary American English (COCA)

The Longman learners' corpus (learner language)

- **المجاميع**

هيئة الجنسية البريطانية (BNC)

كوربوس اللغة الإنجليزية الأمريكية المعاصرة (COCA)

مجموعة المتعلمين من Longman (لغة المتعلم)

**Categories:**

**الفئات :**

<i>Criteria</i>	<i>Type</i>
User	Native vs. learner
Content	monitor vs. sample
Language	mono vs. multilingual
Source	original vs. translation
Time	synchronic vs. diachronic
Remark	plain vs. annotated

Search a word in Google through time!

Word usage .

### Uses & functions:

- 1- Computing: e.g. natural language processing
- 2- Linguistics: e.g. the study of linguistic competence or performance as revealed in naturally occurring data (descriptive grammar) + Critical Discourse Analysis.
- 3- Language teaching: e.g. ESP.

### الاستخدامات والوظائف:

1. الحوسبة: على سبيل المثال معالجة اللغة الطبيعية
2. اللغويات: على سبيل المثال. دراسة الكفاءة أو الأداء اللغوي كما هو موضح في البيانات الطبيعية (القواعد الوصفية) + تحليل الخطاب النقدي.
3. تدريس اللغات: على سبيل المثال ESP

### Corpus linguistics

#### لغويات الاحضار

<i>Criteria</i>	<i>Type</i>
User	Native vs. learner
Content	monitor vs. sample
Language	mono vs. multilingual
Source	original vs. translation
Time	synchronic vs. diachronic
Remark	plain vs. annotated

Search a word in Google through time!

Word usage .

Uses & functions:

2. Computing: e.g. natural language processing
3. Linguistics: e.g. the study of linguistic competence or performance as revealed in naturally occurring data (descriptive grammar) + Critical Discourse Analysis.
4. Language teaching: e.g. ESP.

الاستخدامات والوظائف:

1. الحوسبة: على سبيل المثال معالجة اللغة الطبيعية
2. اللغويات: على سبيل المثال. دراسة الكفاءة أو الأداء اللغوي كما هو موضح في البيانات الطبيعية (القواعد الوصفية) + تحليل الخطاب النقدي.
3. تدريس اللغات على سبيل المثال ESP



## CL and CDA:

CDA = How is X represented in Y source ?

(e.g. Saudi Women in Fox news)

Agent vs. theme

How ?

## L و CDA:

CDA = Y في المصدر X كيف يتم تمثيل

(مثال: سيدات سعوديات في فوكس نيوز)

عامل مقابل الموضوع

كيف؟

Collocations (e.g. look for collocations of Saudi/muslim/women...etc)

Concordance patterns: a list of all of the occurrences of a particular search term in a corpus, presented within the context that they occur in.

Keyword patterns: lexical items that show marked frequency in one text compared to another.

الاستنساخ ( على سبيل المثال ، البحث عن رعايا المملكة العربية السعودية / المسلمة / النساء ... الخ)

أنماط التوافق: قائمة بكافة تكرارات مصطلح بحث معين في مجموعة ، معروضة في السياق الذي تحدث فيه.

أنماط الكلمات الرئيسية: عناصر معجمية تعرض تكرارًا ملحوظًا في نص واحد مقارنةً بنص آخر.

## CL & ELT:

Barlow (2002):

بارلو (2000)

Syllabus design

تصميم المنهج

Materials development

تطوير المواد

Classroom activities

النشاطات الصفية

### Importance:

L2 learners and central/typical uses of a word.

CL & grammar/spelling checkers

Collocation checking

phrase finding - writing classes

Spoken language repository

Covers different topics

Dictionaries vs. corpora

### الأهمية:

المتعلمين والاستخدامات المركزية / النموذجية للكلمة.

القواعد / المدققون الإملائي & CL

فحص التجميع

العثور على العبارة - فصول الكتابة

مستودع اللغة المنطوقة

يغطي مواضيع مختلفة

القواميس مقابل الكومبا

Now we can explore COCA corpus online as an activity.....

Homework (extract X from the COCA corpus)- options are: collocates of X noun .... the most frequently used word ..... concordance lines of a noun of your choice....

الآن يمكننا استكشاف COCA corpus على الإنترنت كنشاط...

الواجبات المنزلية (استخراج X من COCA corpus ) الخيارات هي: رتب من X noun.... الكلمة الأكثر استخداما ... خطوط التوافق من اسم من اختيارك....

### NLP & artificial intelligence

#### البرمجة اللغوية العصبية والذكاء الاصطناعي

- Natural Language Processing
  - Artificial Intelligence
- معالجة اللغة الطبيعية
  - الذكاء الاصطناعي

#### Definition ?

An area of computer science which deals with the study and development of computer systems that are able to perform tasks which require human intelligence.

**تعريف؟** مجال علوم الكمبيوتر الذي يتعامل مع دراسة وتطوير أنظمة الكمبيوتر القادرة على أداء المهام التي تتطلب ذكاء بشري

#### Applications ? التطبيقات:

Visual perception (Face recognition- Cameras /Self-driving cars recognise people - CIA !!)

Decision making (If X happens, do Y - Video games -Tennis)

Problem-solving (photoMath app !!)

Machine translation (google)

Speech recognition systems (Siri !)

Information retrieval (ads)

الإدراك البصري (التعرف على الوجوه - الكاميرات / السيارات ذاتية القيادة التي تتعرف على الأشخاص CIA)

صنع القرار ( إذا حدث X قم بعمل Y - ألعاب الفيديو- تنس )

حل المشكلات ( تطبيق photomatix )

الترجمة الآلية (جوجل)

نظم التعرف على الكلام (سيري)

استرجاع المعلومات (الإعلانات)

## NLP:

### What does it stand for?

Natural language = Any language spoken by humans (e.g. Arabic, English, Zulu !!) Vs. artificial languages such as C++, Java, or Morse code.

البرمجة اللغوية العصبية:

ماذا تعني هذه ؟

اللغة الطبيعية = اي لغة يتحدث بها البشر (على سبيل المثال العربية والانجليزية ، الزولو!) مقابل لغات اصطناعية مثل سي ++ ، جافا ، أو شفرة مورس.

### NLP definition ?

Natural language processing (NLP) is a field of computer science, artificial intelligence, and linguistics concerned with the interactions between computers and human (natural) languages.

تعريف البرمجة اللغوية العصبية ؟

تعتبر معالجه اللغة الطبيعية (البرمجة اللغوية العصبية) مجالاً لعلوم الكمبيوتر والذكاء الاصطناعي واللغويات المعنية بالتفاعلات بين الحواسيب واللغات البشرية (الطبيعية).

### Why NLP ?

#### Uses are countless !

To start with, let's say that you have a company and you receive thousands of emails and comments. How can you analyze them?

You want to find the representation of X in a particular text/speech.

لماذا البرمجة اللغوية العصبية ؟

الاستخدامات لا تعد ولا تحصى..!

بدايةً ، لنفترض أنك تمتلك شركة وتلقى آلاف الرسائل الإلكترونية والتعليقات. كيف يمكنك تحليلها؟  
تريد العثور على تمثيل X في نص / خطاب معين.

## Corpus Linguistics

## اللغويات الأساسية

## SOME INSIGHTS OBTAINABLE FROM CORPORA

... but not maybe all obtainable by us from corpora we have free access to...

Most of these have fairly obvious use for both descriptive linguists and teachers... and maybe learners too (and others in the range of users).

بعض الاحصاءات المستفادة من كوربورا:

ولكن ليس ربما كل ما يمكن الحصول عليه من طرفنا نحن نملك حرية الوصول إليه.....  
معظمها له استخدام واضح إلى حد ما لكل من اللغويين والمعلمين الوصفية ... وربما المتعلمين أيضًا (والبعض الآخر في نطاق المستخدمين).

## General English

Frequencies of individual words across varieties: *certain* and *sure*

Characteristics of varieties and individual authors: frequencies overall; TTRs

Details of meaning of vocabulary items and collocation: qualitative details of synonyms *sad-unhappy*; mutual information for *money* and *flatly*

Homonym and sense frequencies: *lookout*

Lexical grammar: verbs used with *that* clauses

Grammar: uses of *with*

Use of words with a heavy pragmatic dimension: *flipping*, *right*

Lexical phrases: *You know what I mean...*

اللغة الإنجليزية العامة:

ترددات الكلمات الفردية عبر أصناف: مؤكدة ومضمونة

خصائص الأصناف والمؤلفين الفرديين: ترددات عامة؛ TTRs

تفاصيل معاني المفردات والترميز: التفاصيل النوعية للمرادفات الحزينة غير السعيدة؛ المعلومات المتبادلة للحصول على المال وبشكل قاطع

والتردد: المراقبة homonym ترددات

قواعد اللغة المعجمية: الأفعال المستخدمة مع تلك الجمل

القواعد: استخدامات مع

استخدام الكلمات ذات البعد البراغماتي الثقيل: التقليل ، صحيح

عبارات معرفية: أنت تعرف ما أقصد...

- **Translation**

Frequency of translation equivalences

- **Learner Language**

Error and performance analysis information from teacher-made mini-corpora of their learners' language

Ditto from large corpora of learner language

Frequency of types of lexical error

Research on error correctability by dictionaries

- **ترجمة**

تردد معادلات الترجمة

- **لغة المتعلم**

معلومات تحليل الأخطاء والأداء من المصنفات المصغرة المدروسة من لغة المتعلمين، كما سبق من مجموعة كبيرة من لغة المتعلم.

تردد أنواع الخطأ المعجمية بحث عن تصحيح الأخطاء عن طريق القواميس.

## GENERAL TYPES OF CORPUS TASK/ INVESTIGATION

It is possible classify most corpus projects, or generate new ones, as combinations of choices from these main dimensions (for any given language, assumed to be English here):

أنواع عامة من مهمة CORPUS / التحقيق

من الممكن تصنيف معظم مشاريع corpus ، أو توليد مشاريع جديدة ، كمجموعة من الاختيارات من هذه الأبعاد لأي لغة معينة ، يفترض أن تكون الإنجليزية هنا (الرئيسية).

### 1. What kind of language is of interest?

from normal native speaker adults today. Then it could be spoken or written, standard or non-standard, UK or US or..., from everyday language or the specialist register of newspapers or poetry or academic prose or...etc.

from the past. Literary or not...

from foreign language learners

from normal native speaker children

from speakers with language disabilities (e.g. aphasics)

## 1. ما نوع اللغة محل الاهتمام؟

من البالغين اللغة الأم الطبيعية الأصلية اليوم. ثم يمكن أن تكون منطوقة أو مكتوبة ، أو قياسية أو غير قياسية ، أو المملكة المتحدة أو الولايات المتحدة أو ... ، من اللغة اليومية أو سجل متخصص في الصحف أو الشعر أو النثر الأكاديمي أو ... إلخ.

من الماضي. الأدبية أم لا،...

من المتعلمين باللغة الاجنبية من الأطفال العاديين الناطقين باللغة الاصليه ،

من المتكلمين ذوي الإعاقات اللغوية (مثل الملاحم) .

## 2. What level(s) of language are you interested in?

vocabulary/lexis

grammar/syntax

sounds, intonation

spelling, punctuation

text/discourse/rhetorical structure

pragmatics

## 2. ما هو مستوى (مستويات) اللغة التي تهتم بها؟

المفردات / المعجم

قواعد اللغة / تركيب

الأصوات ، التجويد

الهجاء وعلامات الترقيم

النص / الخطاب / الهيكل البلاغي

التداولية

- **Spoken? how natural are speeches, TV etc.?**

Fully natural: observer's paradox and how to be ethical? Permission. Labov's tricks

Records of speakers (and addressees and...)

Transcription issues: what to transcribe and who does it (expert or not)

Random sampling again; problem of accents and dialects

- **تحدث؟ ما مدى طبيعية الخطب والتلفزيون وما إلى ذلك؟**

طبيعي تماما: مفارقة المراقب وكيف تكون أخلاقية؟ الإذن. حيل لابوف

سجلات المتكلمين (والعناوين و...)

قضايا النسخ: ما نسخها ومن يفعل ذلك (خبر أم لا)

أخذ العينات عشوائية مرة أخرى. (مشكلة اللهجات واللهجات)

- **Analysis - how to extract useful information automatically?**

frequency and its derivatives:

range: over text types

richness of vocab: TTR

collocational strength: mi and t-score/z score

how to relate *go*, *goes* and *went*? lemmatisation

concordance: the problem of large numbers. Qualitative into quantitative

how to distinguish *right* from *right*: pos and other annotation/tagging

- **تحليل - كيفية استخراج المعلومات المفيدة تلقائياً؟**

التردد ومشتقاته:

النطاق: عبر أنواع النص

ثراء من المفردات: TTR

القوة المهنية mi و t-score / z score



كيفية ربط الذهاب ، وغادرت؟ lemmatisation

التوافق: مشكلة الأعداد الكبيرة. النوعية إلى الكمية

كيفية تمييز الحق من اليمين pos: وغيرها من التعليقات / التمييز

- **how to sort and select from a KWIC listing?**

Accessibility to general users – cost, computers etc.

- **KWIC كيفية فرز واختيار من قائمة ؟**

إمكانية الوصول إلى المستخدمين بشكل عام - التكلفة ، أجهزة الكمبيوتر إلخ

### **3. What is the purpose of the investigation? More language research, or more pedagogical?**

- to describe an aspect of language or compare different styles, authors etc. I.e. more exploratory research.
- to check on a proposed 'rule' or past finding or a theory-based prediction in some area of language study. I.e. more hypothesis testing research.
- to test out a parser that some computational linguists have designed
- to help create language syllabuses or teaching materials
- to help evaluate syllabuses or teaching materials
- to use or evaluate corpus work as a class task (i.e. a form of CALL)
- to help write a dictionary or grammar book
- to help evaluate a dictionary or grammar book
- etc.

### **3. ما هو الغرض من التحقيق؟ مزيد من البحوث اللغوية ، أو أكثر تربوية؟**

- لوصف جانب من جوانب اللغة أو مقارنة الأنماط المختلفة ، المؤلفين إلخ. مزيد من البحوث الاستكشافية.
- للتحقق من "القاعدة" المقترحة أو النتيجة السابقة أو التنبؤ القائم على النظرية في بعض مجالات دراسة اللغة. أي مزيد من الفحوصات اختبار البحوث.
- لاختبار محلل أن بعض اللغويين الحاسوبيين قد صمموا
- للمساعدة في إنشاء مناهج اللغة أو مواد التدريس
- للمساعدة في تقييم المناهج الدراسية أو المواد التعليمية
- للمساعدة في كتابة قاموس أو كتاب قواعد استخدام أو تقييم العمل كمهمة طبقية (مثل نموذج CALL)

- للمساعدة في تقييم قاموس أو قواعد اللغة
- ... إلخ

#### 4. What kind of corpus information is needed?

- more concordance-type information - examples of occurrences of things in context to analyse. I.e. qualitative
- more frequency information about words or whatever. I.e. quantitative
- Most of the combinations implied above are possible to some extent with existing corpora. However, they are not all available to us here.

#### 4. ما هو نوع المعلومات الأساسية المطلوبة؟

- مزيد من المعلومات من نوع التوافق - أمثلة على حدوث أشياء في سياق لتحليلها (نوعي)
- معلومات أكثر تردد حول الكلمات أو أيا كان (كمي)
- معظم التوليفات المتضمنة أعلاه ممكنة إلى حد ما مع الكومبا الموجودة. ومع ذلك ، فهم ليسوا جميعًا متاحين لنا هنا.

- BNC and suggested tasks
- Corpus Linguistics.
- What is BNC?

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

- المهام المقترحة
- لغويات الجسم.
- ما هو الفن الوسيطي ؟

- Suggested Tasks.

### SOME SUGGESTED CORPUS TASKS

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. 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. Similarly, corpus use, like introspection or administering tests or questionnaires to subjects, is not in itself usually a project in itself. Rather it is a means to carry out some project in language description, language teaching or whatever. The bulk of the project has to come from the user's prior knowledge of linguistics, teaching etc.

- المهام المقترحة:

#### بعض المهام المقترحة:

الشيء المهم الذي يجب أن ندركه بالطبع هو أن المجاميع ومحركات البحث تشكل في المقام الأول أدوات أو طرق بحث ، بدلاً من مجالات التحقيق في حد ذاتها.

قليل من الناس يدرسون المطارق بدلاً من ذلك يستخدمونها لأداء المهام التي يفكرون بها ، مثل بناء خزانة ، والتي لديها أبعاد بعيدة عن المطارق التي يتعين على المستخدم جلب الكثير من الخبرة المنفصلة إليها. وبالمثل ، فإن استخدام الجسم ، مثل الاستبطان أو إجراء اختبارات أو استبيانات للموضوعات ، ليس في حد ذاته مشروعًا بحد ذاته. بل هو وسيلة لتنفيذ بعض المشاريع في وصف اللغة ، أو تدريس اللغة أو أي شيء. يجب أن يأتي الجزء الأكبر من المشروع من معرفة المستخدم السابقة باللغويات والتدريس ، إلخ.

Therefore in choosing a task you have to think what linguistics you know most about already, and choose a task accordingly. Those suggested below are mostly descriptive linguistic (vocabulary and grammar mainly), or involve some pedagogical evaluation or authoring with a corpus element.

لذلك عند اختيار مهمة ، عليك أن تفكر في ما هي اللغويات التي تعرفها أكثر من غيرها ، وأن تختار مهمة وفقًا لذلك. تلك المقترحة أدناه هي لغوية وصفية في الغالب (المفردات وقواعد اللغة بشكل رئيسي) ، أو تنطوي على بعض التقييم التربوي أو التأليف مع عنصر corpus .

They should be do-able with the corpus and concordancing resources you can access this year, though I cannot guarantee anything as what is available changes by the minute. 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).

يجب أن يكونوا قادرين على التعامل مع الموارد والموارد التوافقية التي يمكنك الوصول إليها هذا العام ، على الرغم من أنه لا يمكنني ضمان أي شيء على أنه التغييرات المتاحة في كل دقيقة. يمكنك بالطبع التفكير في مشاريعك الخاصة وفقًا لمصلحتك الخاصة ، والتواصل مع الدورات الأخرى التي قد تقوم بها (نظرًا لأن أي دورة تقربها في قسم اللغة واللغويات من المحتمل أن يكون لها بعد أساسي).

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. Often the instructions you can enter in the search engine will not produce all and only the information you want.

ما يلي لم يتم العمل بها بالكامل ، وبأي ترتيب معين. عليك أن تقرر ما إذا كانت المجموعة ذات الصلة متوفرة ، وبالطبع احصل على محرك البحث للبحث عن معلومات مفيدة. غالبًا ما لا ينتج عن التعليمات التي يمكنك إدخالها في محرك البحث جميع المعلومات التي تريدها فقط.

The trick is to get as close as possible, and then sort through the output by hand for what you need..... and interpret it. Make good use of your intuitions as a teacher and/or descriptive linguist!

الخدعة هي أن تقترب قدر الإمكان ، ومن ثم قم بفرز الناتج باليد ما تحتاج إليه ..... وتفسيره. استفد من حدسك كمعلم و / أو لغوي وصفي.

- **How do 'synonyms' differ?**

In class we look briefly at the 'synonyms' *sad* and *unhappy*. You could look at another pair of 'synonyms' like any of those below.

- **كيف تختلف "المرادفات"؟**

في الصف نبحث بإيجاز في "المرادفات" المحزنة وغير السعيدة. يمكنك النظر إلى زوج آخر من "المرادفات" مثل أي من تلك الموضحة 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 ، Leech ، Zgusta إلخ ، حتى يكون هناك اختلاف في أنواع الطرق التي قد تختلف بها بمهارة.

Get concordance output from a suitable set of texts. Give an account of the similarities and differences between the chosen synonyms. How much of your analysis is from the corpus, how much from introspection prompted by the corpus information?

الحصول على إخراج التوافق من مجموعة مناسبة من النصوص. أعط حساباً عن أوجه التشابه والاختلاف بين المرادفات المختارة. كم من التحليل الخاص بك هو من الجسم ، كم من الاستبطان مدفوعاً بمعلومات الجسم؟

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.

يمكنك أيضاً الإشارة إلى الإدخالات في قواميس المرادفات التي لا تدرجها فقط بل تتضمن "مقالات مرادفة" تحاول شرح الاختلافات (مثل قاموس المترادفات في Webster ، أو دليل Cassell الحديث للمترادفات المتوفر في مكتبة Colchester العامة ، Trinity ساحة ، وسط المدينة) كمصدر للأفكار حول ما هي الاختلافات وشيئاً ما ينتقده.

See also error and usage books like Alexander: Right Word Wrong Word and Heaton and Turton: Dictionary of Common Errors. These are for foreign learners. You could use them as an aid to your own analysis or do a critique - do they mislead?

راجع أيضاً كتب الأخطاء والاستخدام مثل Alexander: Right Word Wrong Word و Heaton و Turton: Dictionary of Common Errors. هذه للمتعلمين الأجانب. يمكنك استخدامها كمساعد لتحليلك الخاص أو القيام بنقد - هل هم مضللون؟

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 مهتمًا بتدريس اللغة فقد ترغب في التفكير في كيفية اختيار وتهيئة خطوط corpus التي تجدها لعمل مهمة تفاضل مرادف فعالة لبعض المتعلمين المحددين الذين تضعهم في اعتبارك.

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.

يمكنك أيضًا التفكير في المرادفات ضمن مجموعة محددة من اللغة الإنجليزية (على سبيل المثال ، الكتابة الأكاديمية). بدلاً من اختيارها عمومًا ، من خلال اختيار مجموعة داخل BNC على سبيل المثال.

- **Distinguishing 'confusibles'.**

There are some well-known 'confusibles' which are similar in sound as well as meaning (cf dictionaries by Room). Some of these may also have proscriptions associated with them. See also the points made about the last task.

- " تمييز confusables "

هناك بعض confusibles المعروفة والتي هي مشابهة في الصوت وكذلك المعنى (راجع القواميس). بعض هذه قد يكون لها أيضًا proscriptions المرتبطة بها. انظر أيضًا النقاط التي تم القيام بها 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 أن يساعد بحث كوربوس أحدًا في تحديد ما هي الاختلافات بالفعل ، وما إذا كانت العبارات في القواميس (على سبيل المثال ، ملاحظات الاستخدام في قاموس لونغمان للغة الإنجليزية المعاصرة أو قاموس أكسفورد المتقدم للمتعلمين) حول كيفية استخدامها صحيحة..

production, produce, product      إنتاج ، إنتاج ، منتج

continual, continuous      مستمر ، مستمر

comprise, compose, consist of, constitute, include

تتألف ، تتكون ، تتكون ، وتشكل ، وتشمل.

## ▪ Future time expression over the years.

A reported feature of the recent history of English has been the rise of the use of *BE going to..* to express future time.

### ▪ التعبير الزمني في المستقبل علي مر السنين:

وقد تم الإبلاغ عن سمه من سمات التاريخ الحديث للغة الانجليزية وكان ارتفاع استخدام الذهاب إلى.. للتعبير عن الوقت في المستقبل.

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').

هل يمكن ان ننظر في بعض النصوص من اللغة الانجليزية الحالية ، والكتاب الأقدم (مثل ديكنز) ، وحتى شكسبير لمعرفة عدد المرات التي يحدث فيها هذا التعبير ، وكيف غالبا ما يبدو ان تستخدم في المستقبل بمعنى (وليس فقط حرفيا يعني ' التحرك نحو..

What characters use it? (Innovations often start socially 'from below'). And perhaps you could look at one or more other means of conveying future time in the same texts. Note you would need to try and match up styles of text as far as possible across the years.

ما هي الأحرف التي تستخدمها ؟ (غالبا ما تبدأ الابتكارات اجتماعيا من الأسفل). وربما يمكن ان ننظر إلى واحد أو أكثر من وسيله أخرى لنقل الوقت في المستقبل في نفس النصوص. ملاحظه لك تحتاج إلى محاولة والمباراة حتى أنماط النص بقدر الإمكان عبر السنوات.

Refer also to Quirk et al. or Leech on time and tenses in English.

Frequency and a 'lexical syllabus' for learners

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.

الاشاره أيضا إلى الهوس وآخرون أو علقه في الوقت والازمنه باللغة الانجليزية.

تردد و ' المنهج المعجمي ' لمقرري المناهج التعليمية وكثيرا ما حاولت السيطرة علي إدخال البنود vocaloid في دوره اللغة ، والمعيار الأكثر شعبية وكان التردد. اي ان الدورة التدريبية الجديدة تقدم تقريبا من أجل التردد المتناقص في اللغة المستهدفة ، استنادا إلى بعض العدد.

The most popular count relied on in EFL from the 40s onwards was Michael West's General Service List of English Words. 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. See books by Willis and Lewis, and the COBUILD English Course.

الاعتماد الأكثر شعبيه المعتمدة علي في المنتج من الاربعينات فصاعدا كان مايكل ويست الخدمات العامة في قائمه الكلمات الانجليزيه. في الاونه الاخير هذه الفكرة قد حصلت علي عقد جديد من الحياة تحت راية "المنهج المعجمي" ، واليوم لدينا تعول علي أساس أكبر بكثير من الغرب في 5,000,000 أو نحو ذلك. انظر الكتب من قبل ويليس ولويس ، والدورة الانجليزيه المشتركة

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. You would have to sample the items and check their frequency and produce a profile. Or scan them and use the Compleat Lexical Tutor online. If not selected by frequency, by what criteria then?

هل يمكن ان تاخذ كتاب الدورة التي تسرد الكلمات الجديدة في ذلك ، أو المنهج الذي يوفر قوائم الكلمات التي يجب ان تكون معروفه من قبل مستويات مختلفه ، ونري إلى اي مدي يبدو ان اختيار والدرجات في اتفاق مع التردد. سيكون لديك لعينه العناصر والتحقق من تواترها وإنتاج ملف التعريف. أو مسحها واستخدام المعلم المعجمي علي الإنترنت. إذا لم يتم تحديدها بواسطة التردد ، فما هي المعايير بعد ذلك ؟



## Lecture 12

- NLP
- Natural Language Processing
- What is NLP?

- Natural Language Processing (NLP)

Computers use (analyze, understand, generate) natural language

A somewhat applied field

Computational Linguistics (CL)

Computational aspects of the human language faculty

More theoretical

- Nlp

- معالجه اللغة الطبيعية
- ما هي البرمجة اللغوية العصبية ؟

- معالجه اللغة الطبيعية (البرمجة اللغوية العصبية) :

- تستخدم أجهزة الكمبيوتر (تحليل وفهم وتوليد) اللغة الطبيعية.
- حقل تطبيقي إلى حد ما
- اللغويات الحاسوبية CL
- الجوانب الحسابية للكلية اللغة البشرية
- أكثر نظرية

- Why Study NLP?

- Human language interesting & challenging
- NLP offers insights into language
- Language is the medium of the web
- Interdisciplinary: Ling, CS, psych, math
- Help in communication
- With computers (ASR, TTS)
- With other humans (MT)
- Ambitious yet practical

- لماذا تدرس البرمجة اللغوية العصبية ؟

- لغة إنسانية شيقة وصعبة
- تقدم البرمجة اللغوية العصبية رؤى في اللغة
- اللغة هي وسيط لشبكة الإنترنت
- متعدد التخصصات: لينغ ، CS ، نفسي ، رياضيات
- تساعد في التواصل:
- مع أجهزة الكمبيوتر ( TTS ، ASR )
- MT مع البشر الآخرين
- طموح بعد عملي

- Goals of NLP :

- Scientific Goal:

*Identify the computational machinery needed for an agent to exhibit various forms of linguistic behavior*

- Engineering Goal:

*Design, implement, and test systems that process natural languages for practical applications*

- أهداف البرمجة اللغوية العصبية:

- الهدف العلمي:

تحديد الآلية الحسابية اللازمة للوكيل لعرض أشكال مختلفة من السلوك اللغوي.

- الهدف الهندسي:

تصميم وتنفيذ واختبار النظم التي تعالج اللغات الطبيعية للتطبيقات العملية.

- Applications :

- speech processing: get flight information or book a hotel over the phone
- information extraction: discover names of people and events they participate in, from a document
- machine translation: translate a document from one human language into another
- question answering: find answers to natural language questions in a text collection or database
- summarization: generate a short biography of Noam Chomsky from one or more news articles

- التطبيقات:

- معالجة الكلام: الحصول على معلومات عن الرحلات الجوية أو حجز فندق عبر الهاتف.
- استخراج المعلومات: اكتشاف أسماء الأشخاص والأحداث التي يشاركون فيها.
- الترجمة الآلية: ترجمة وثيقة من لغة إنسانية إلى لغة أخرى.
- الإجابة على الأسئلة: العثور على إجابات لأسئلة اللغة الطبيعية في مجموعة نصية أو قاعدة بيانات
- قاعده البيانات: إنشاء سيرة ذاتية قصيرة من نصوص تشومسكي من واحد أو أكثر من المقالات الإخبارية

- General Themes :

- Ambiguity of Language
- Language as a formal system
- Rule-based vs. Statistical Methods
- The need for efficiency

- المواضيع العامة:

- غموض اللغة
- اللغة كنظام رسمي
- القواعد المعتمدة مقابل الأساليب الإحصائية
- الحاجة إلى الكفاءة

- Ambiguity of language :

Phonetic [raɪt] = write, right, rite

Lexical can = noun, verb, modal

Structural I saw the man with the telescope

Semantic dish = physical plate, menu item

All of these make NLP difficult

- غموض اللغة:

الغموض الصوتي: [raɪt] = write, right, rite

الغموض المعجمي: can = noun, verb, modal

غموض التراكيب: I saw the man with the telescope

الغموض السيميائي "الدلالي": dish = physical plate, menu item

كل هذا يجعل البرمجة اللغوية العصبية صعبة.

- **Language as a formal system :**
  - We can treat parts of language formally

Language = a set of acceptable strings

Define a model to recognize/generate language

- Works for different levels of language (phonology, morphology, etc.)

Can use finite-state automata, context-free grammars, etc. to represent language

- **اللغة كنظام رسمي:**

- يمكننا التعامل مع أجزاء من اللغة بشكل رسمي

اللغة = مجموعه من السلاسل المقبولة

تحديد نموذج للتعرف على اللغة أو توليدها.

- يعمل على مستويات مختلفه من اللغة (علم الصوتيات ، علم التشكل ، الخ)

يمكن استخدام الآليات ذات الحالة المحدودة ، القواعد النحوية الخالية من السياق ، وما إلى ذلك لتمثيل اللغة.

- **Rule-based & Statistical Methods:**

Theoretical linguistics captures abstract properties of language

NLP can more or less follow theoretical insights

- Rule-based: model system with linguistic rules
- Statistical: model system with probabilities of what normally happens

Hybrid models combine the two

- **الأساليب القائمة علي القواعد والإحصاءات:**

اللغويات النظرية تلتقط الخصائص المجردة للغة.

البرمجة اللغوية العصبية يمكن ان تتبع أكثر أو اقل رؤى النظرية.

- القاعدة قائمة على: نظام نمذجي بقواعد لغوية
- الإحصاء: نظام نمذجي مع احتمالات ما يحدث عادة

نماذج مختلطة تجمع بين الاثنين.

- **The need for efficiency :**

- Simply writing down linguistic insights isn't sufficient to have a working system
- Programs need to run in real-time, i.e., be efficient
- There are thousands of grammar rules which might be applied to a sentence
- Use insights from computer science
- To find the best parse, use chart parsing, a form of dynamic programming

- **الحاجة إلى الكفاءة:**

- لا يكفي مجرد تدوين الرؤى اللغوية أن يكون لديك نظام عمل
- يجب تشغيل البرامج في الوقت الفعلي ، أي أن تكون فعالاً
- هناك الآلاف من القواعد النحوية التي يمكن تطبيقها على الجملة
- استخدم رؤى من علوم الكمبيوتر
- للعثور على أفضل تحليل ، استخدم تحليل الرسم البياني ، وهو شكل من أشكال البرمجة الديناميكية.

Exploring corpora

استكشاف المجاميع

The Corpus of Contemporary American English

: <https://corpus.byu.edu/coca/>

مجاميع اللغة الإنجليزية الأمريكية المعاصرة

: <https://corpus.byu.edu/coca/>