**Introduction to
L2 Vocabulary Acquisition
& Learning:
Lecture 10**

Yukio Tono

Fall 2015

VOCABULARY AS DATA

2016/1/5   2

Vocabulary as data for learning

- In this talk, we look at what problems researchers face when they confront large amounts of vocabulary in the form of various kinds of data and try to make systematically meaningful statements about how often and where words occur, and in what sorts of envieonments.

2016/1/5   3

TASK

- Somebody offers you a list of what they claim are 'the 1,000 most frequent words in language X'. What questions would you want to ask them about their list and the way in which it was created?

2016/1/5   4

Possible questions about any frequency list (1)

1. What size corpus was used to get the frequency count?
2. Was the corpus written language, spoken, or both?
3. Did the corpus cover a wide range of text-types, topics, registers, situations, etc.?
4. Did the frequency count bunch word-forms together under single entries, or did it separate them? In other words, are 'actual' and 'actually' treated as the same item or as different?

2016/1/5   5

Possible questions about any frequency list (2)

5. How long ago was the corpus assembled? Does it contain up-to-date information?
6. Does the count take into consideration very frequent multi-words items? For example, in English, we might wish to know whether '*get up*' is more frequent than '*get on*,' or '*get away with*'.
7. Does the count tell us about frequency of meanings?

2016/1/5   6

M. West (1953), General Service List

FAT

fat, adj.			
	(1) A fat man, sheep		
		Fat meat	61%
	(2) (=*thick*)		
		Plant with fat leaves	14%
	(3) (*figurative*)		
		Fat pastures	3%
fat, n.	*Cooking fat, beef fat, Oils and fats*		18%

TASK

- Look at the entry 'save'. Can you make approximate guesses as to which meanings of the word will be most frequent? What factors influence your decisions? Check your guesses against West's percentages.

SAVE (v.)

(1) Save a person from danger
Just saved himself in time
I saved my books from the fire. (%)
(2) (protect against possible danger or loss)
A railing to save people from falling over.
God save the King. (%)
(3) (economize)
Save time, trouble, cost
A great saving of time (n.) (%)
(4) (store up against the future)
Save money
Save up (money)
My savings (n.)
The savings-bank (n.) (%)

SAVE (v.)

(1) Save a person from danger
Just saved himself in time
I saved my books from the fire. (17 %)
(2) (protect against possible danger or loss)
A railing to save people from falling over.
God save the King. (24 %)
(3) (economize)
Save time, trouble, cost
A great saving of time (n.) (9 %)
(4) (store up against the future)
Save money
Save up (money)
My savings (n.)
The savings-bank (n.) (26 %)

RANGE

- A word (or word-form) may be quite frequent, but a majority, or even all, of its occurrences might be in just one or two texts, in which case, although its *frequency* might look significant, its *range* might be quite small.
- The useful words for the learner are those words which (a) are **frequent**, and (b) have a fairly **wide range**, that is, those which occur across a wide variety of texts.
- Sometimes this measure is called **dispersion**.

TASK

- Compare the two extracts below. The first is a transcript of two people talking while they are trying to adjust a door-closure to prevent it from slamming. The second is a written set of instructions accompanying a picture of the door-closure mechanism. What are the main differences in vocabulary between the two?

Extract 1

Mike: erm … that's too much now, too slow, I'll have to take it back a bit … there
Donna: oh, there's two
Mike: this one controls the speed
Donna: I don't mind if it doesn't shut properly at all
Mike: no, it'll shut, it'll shut … so … try it now
Donna: yeah, that's it
Mike: it's better than it was
Donna: it'll do

Extract 2

- The outer screwing mechanism (A) controls the snapshut operation and should be adjusted clockwise to reach the slowest speed while still affecting proper closure. The inner screw (B) adjusts clockwise to control the overall closing speed and may be set to the desired rate of closure. Replace the protective cover and tighten the retaining screws (C and D).

Lexical density

- Both extracts are just over fifty words long.
- But it is noticeable that the written text seems 'heavier' in vocabulary than the spoken one.
- The spoken text is an example of 'language-in-action', that is, people are using language as an accompaniment to the action they are engaged in, and the feeling of lightness or heaviness of vocabulary is what Ure (1971) calls '**lexical density**'.
- The lexical density of a text can be measured by counting the total words in a text and then counting the *lexical* words, that is, the content words, excluding the grammar or function words, and calculating the lexical words as a percentage of the total words;
- The higher the percentage, the higher the lexical density.
  Spoken   51 words   15 lexical words   29.4%
  Written   53   33   62.2%

Lexical variation

- Lexical variation takes as its starting point the distinction between token and type. If a text is 100 words long, it is said to contain 100 *tokens*, but many of these tokens may be repeated within the text and this may give us a considerably lower total of *types*. The ratio between tokens and types is called **type/token ratio (TTR)**.

TASK

- Do a type/token ratio for this short text:

They came yesterday to sort out some problem with the phones. The phone in my office has somehow got crossed with the secretary's and they just couldn't work out how the wiring system was, since it was so archaic. They have not solved the problem yet and have to come again tomorrow.

Some observation

- Certain decisions have to be made in counting which might affect the result:
  - Are 'came' and 'come' the same type?
  - Is the 'out' of 'sort out' the same 'out' as in 'work out'
- Consistency is all important.
  - It is probably sensible for pedagogic ends to treat inflected forms of a word as the same type.
- Lexical variation counts do give us a rough measure of how many new items are introduced into a text as it unfolds.

SELECTING WHAT TO TEACH

2016/1/5   19

Selecting what to teach

- Criteria for selection of vocabulary:
  - Frequency
  - Range
  - Learnability
  - Learner's needs

2016/1/5   20

Frequency

- Most general course books take for granted that the most frequent words in the language must be presented in the early stages of language teaching, and this is a sensible decision.
- However, frequency lists need to be approached with caution and careful analysis, as suggested in the last lesson.

2016/1/5   21

Some cautions of using frequency

- Major corpora such as LOB and the Birmingham corpus do not agree either in every detail, though taken as a whole, there is substantial agreement on the high frequency of the most common words.
- In frequency lists, words from the same subject area are not necessarily of the same frequency and words of similar frequency come from wide different subject area.

2016/1/5   22

Empty vs. available words

- The most frequent words are usually the most informationally *empty* words and yet the learner needs to communicate and understand messages with considerable content right from the very beginning.
- *Available* (=most disposable) words:
  - Available words are known in the sense that they come to mind rapidly when the situation calls for them. (Richards 1974)

2016/1/5   23

Learnability

- Words may be easy or difficult for a variety of reasons other than frequency.
1. Words may present spelling difficulties. (e.g. occurrence, parallel)
2. Words may present phonological difficulties (e.g. thrive, crisps)
3. The syntactic properties of words often make them difficult.
4. Words may be perceived as very close in meaning by the learner, and therefore difficult to separate one from another.
5. Words may be 'false friends' (e.g. *Aktuellt* in Swedish, *actuellement* in French, do not mean the same as English *actually*.)
6. Learners may be unable to relate the meaning of a word to their world experience or to their culture. (e.g. deacon)

2016/1/5   24

TASK

- Write down five words in any foreign language you know, and then consider whether the words were easy or difficult for you to learn, and for what reasons.

Learners' needs

- Predicting what learners will need in the way of vocabulary is important in selecting what to teach.
- Equally important is 'creating a sense of need for a word', and recognizing that learner perceptions of need may conflict with the teacher's perceptions.

Predicting learners' vocabulary needs

- Which words must students know in order to talk about people, things, and events in the place where they study and live?
- Which words must the student know in order to respond to routine directions and commands?
- Which words are required for certain classroom experiences (describing, comparing, and classifying various animals, or having imaginary conversations with speakers of English, or writing letters)?
- Which words are needed in connection with the students' particular academic interests?

ORGANIZING VOCABULARY

Organizing vocabulary

- Topic
- Meaning
- Form
- Context relations: register & discourse

Topic

- The use of topic as a framework for vocabulary presentation is very common in materials. However, topic may be a problematic framework for several reasons:
  - It is difficult to actually define a 'topic.'
  - Which topics are going to be most useful for learners anyway?
  - It is difficult to predict exactly what words are the most frequent or useful within any given topic.
  - Topics relate more readily to people's experience than semantic or formal categories may do.

TASK

- Make a list of 10 topic-headings for a ten-week, one-hour-a-week vocabulary development course with you feel would be most useful for a group of lower-intermediate learners of English.
- Compare your results with your classmates and discuss the differences. What factors determined your choice?

2016/1/5   31