



WHAT RESEARCH CAN TEACH US ABOUT TEACHING SECOND LANGUAGE VOCABULARY

Vedran Dronjic
vedran.dronjic@nau.edu

The complexity of lexical knowledge

Knowing a word means knowing a lot:

- **Receptive** or **productive** knowledge
- **FORM**: sound segments, stress, lexical tone, spelling, morphological structure
- **MEANING**: sense, reference, denotation, connotation, polysemy
- **“WORD GRAMMAR”**: syntactic environments a word appears in, morphology it triggers in other words
- **COLLOCATION**: *fast food – ~~quick~~ food, fast shower – quick shower, fast train – ~~quick~~ train*
- **NETWORK KNOWLEDGE**: meaning neighbours, sound neighbours, morphological neighbours, synonyms, antonyms, superordinates, subordinates, etc.
- **SOCIOLINGUISTIC & PRAGMATIC KNOWLEDGE**: When and how can the word be used?
- **Clearly, there can be no one correct way of teaching or learning all this! We begin by matching a form and a meaning.**

Vocabulary size

How do we count words?

- **SURFACE WORD FORMS:** *play, plays, played, playing* = 4
- **LEXICAL ENTRIES:** *play, plays, played, playing* = 1? 2 (verb, noun)? More (*play cards* vs. *play someone*)?
- **WORD FAMILIES:** *play, plays, player, unplayable, playful, playfulness ...* = 1

But: *depart, department* = 2

Also: *horseplay, playtime* = 2

(Un)known vocabulary and text coverage

- 2 k families → 86% of a text is understood (a good start)

Reading with a vocabulary of 2,000 families (86% known words)

If _____ planting rates are _____ with planting _____ satisfied in each _____ and the forest milled at the earliest, the _____ wood supplies could further increase to about 36 million _____ meters _____ in the period 2001-2015.

(Nation, 1990)

(Un)known vocabulary and text coverage

- 2 k families → 86% of a text is understood (a good start, but not enough!)
- Diminishing returns (% of the British National Corpus, excluding names):
 - First 1 k families: 77.96%
 - Second 1 k families: 8.10%
 - Third 1 k families: 4.36%
 - Fourth 1 k families: 1.77%
- 3 k – 4 k families → 95% of a text (15 unknown words on a page of 300; heavy dictionary use; choppy reading experience)
- 9 k families → 98% of a text (6 unknown words per page)
- 7 k families → 98% of a conversation
- 20 k → rank-and-file L1 English-speaking university graduate

Some questions to consider

Should we actively teach vocabulary and why?

- Vocabulary is the **best predictor of comprehension** (at least 50% of reading and listening scores).
 - It will not “teach itself”.
 - Learners don’t know enough of it:
 - English majors in China: 4,000 families after 2,400 hours of instruction
 - Japanese university students: 2,300 families after 1,200 hours
 - Israeli high school graduates: 3,500 families after 1,500 hours
- (Laufer, 2000)

Do students in a typical classroom learn useful words?

- Naturalistic vocabulary learning is largely driven by **frequency of encounter**.
- In L2 learning, frequency is distorted by curriculum, teacher decisions, learning situations, and learning materials; typically not geared toward vocabulary learning
- Learners in classrooms without a special focus on vocabulary frequently do not fully learn the first 2000 families (Cobb & Horst, 2011).
- When there is instruction, it is often **haphazard**: *cummerbund*, *grungy* (neither in the first 16,000 families; Horst, 2014).

Mid-frequency vocabulary (Schmitt & Schmitt, 2014)

- Despite its tremendous usefulness, neither published textbooks nor teachers do a good job of teaching it.
- 3,001–4,000: *academic, consist, exploit, rapid, vocabulary*
- 4,001–5,000: *agricultural, contemporary, dense, insight, particle*
- 5,001–6,000: *cumulative, default, penguin, rigorous, schoolchildren*
- 6,001–7,000: *axis, comprehension, peripheral, sinister, taper*
- 7,001–8,000: *authentic, conversely, latitude, mediation, undergraduate*
- 8,001–9,000: *anthropology, fruitful, hypothesis, semester, virulent*
- General-use academic vocabulary tends to be mid-frequency.

Can students acquire an adequate vocabulary by doing large amounts of reading for pleasure and guessing unknown words from context?

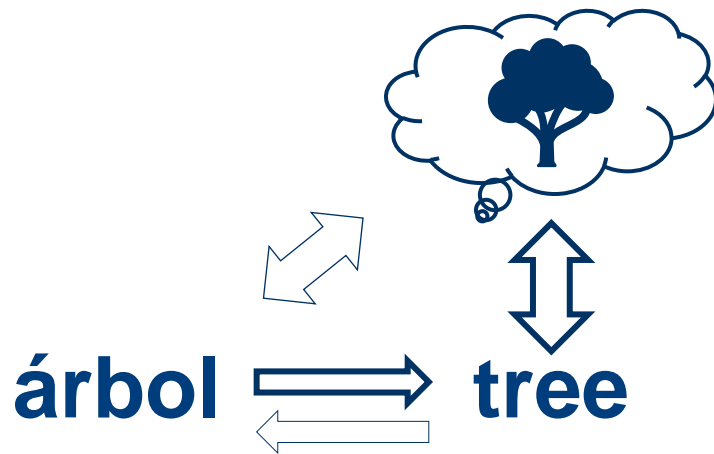
- No.
- At least **6 encounters** for basic form – meaning learning
- Words outside the first 2,000 families are not encountered often enough (Cobb, 2007).
- Guessing is feasible when **1 word out of 20** is not known.
- For a learner who knows 2,000 families, the ratio is 1 in 10.
- When **context supports guessing**, learners don't learn.
- When reading for pleasure, 1 new word is learned per 1,000 words that are read.

Should students use monolingual or bilingual dictionaries?

- What are some arguments in favour of a monolingual dictionary?
- To use a monolingual dictionary, you need a basic vocabulary. Most learners don't have it.
- A well-designed study with 700 Polish learners at all proficiency levels (Lew, 2004):
- What type of dictionary is the best at supporting reading comprehension?
- bilingual > bilingualized > monolingual (across proficiency levels)

Should translation be used when teaching vocabulary?

- A beginning English-speaking learner of Spanish:



- Translation equivalents are **activated automatically** in normal processing.
- Problem with translation: **partial overlap in meaning** (“a plastic glass”)
- **Cognates** (*mother* – *madre*): easier to learn, harder to forget, boost test scores
- Instruction may be needed for **cross-script cognates** (e.g., *інфляція* – *inflation*) ...
- ... and for tricky-to-recognize cognates: *pioggia* – *pluja*; *cane* – *ci*
- **False friends** (related or not): *embarrassed* – *embarazada*; *burro* in Spanish & Italian

Is rote memorization a good way to learn vocabulary?

- It's often reviled as boring and uncreative. It's not “critical thinking”.
- Critical thinking is impossible without a memorized bank of knowledge!
- Learning with **pictures** or through **mnemonic** techniques (*payaso* “clown” – sounds like “pie” – imagine a clown with pie smashed in his face) leads to **inefficient processing**
- Memorization is the single best way of learning form-meaning connections.
- It must be accompanied by **massive input** and **practice!**

Is it a good idea to teach new words in lists of items of the same kind (e.g., *apple, pear, peach, plum*)?

- **Cross-labelling** (🍏 = *pear*) becomes a serious problem.
- ~~*Table, chair, loveseat, ottoman, couch, footstool, sideboard*~~
- *Table, magazine, fruit, mug, (eye)glasses, tablet*: much better
- Best if parts of speech do not overlap (*beach, sunny, swim*)

Are certain kinds of words easier to learn?

- Yes.
- **Concrete** (*leg*) easier than **abstract** (*menace*)
- **Cognate** (*mother - madre*) easier than **non-cognate** (*clock - reloj*)
- **General** (*fatherly*) easier than **register-specific** (*paternal*)
- **Non-idiomatic** (*decide*) easier than **idiomatic** (*make up one's mind*)
- **Few meanings** (*papaya*) easier than **many meanings** (*set*)

Best practices

As a good starting point, see Dronjic (2019), CATESOL Journal (open access)

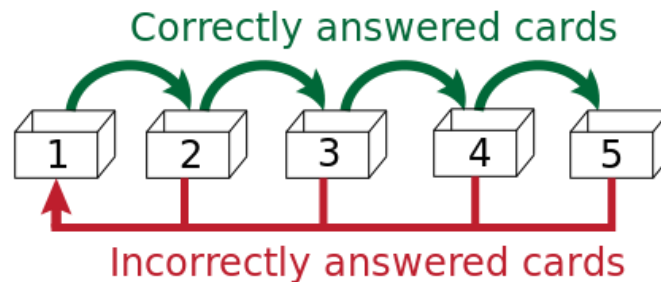
<https://files.eric.ed.gov/fulltext/EJ1238807.pdf>

What to do

- Know your learners' **vocabulary size**.
 - Vocabulary Size Test ([VST](#))
 - Vocabulary Levels Test ([VLT](#)) ([updated VLT](#))
 - More tests [here](#) and [here](#)
- First goal: Systematically **press on with the first 3,000** word families
 - Download [frequency lists](#) and use them to guide teaching
 - Run readings through a [vocabulary profiler](#) and “adjust up or down”
- Teach **mid-frequency** vocabulary (up to 9,000 word families).
- Systematically teach **academic vocabulary** if it's relevant for your students.
- Ignore **low-frequency** vocabulary or address it very quickly and move on (Nation, 2013).

What to do

- **Recycle** the vocabulary you teach. Repeat a few times on the same day (with spacing), then repeat on a few subsequent occasions (slow-wave sleep leads to semantic memory consolidation). **Have regular vocabulary review tests.**
- Encourage systematic autonomous study: e.g., vocabulary diaries, **Leitner box**



- Both the L1 → L2 and L2 → L1 directions!
- There are flashcard [apps](#) that simulate the Leitner box (for computers and mobile devices).

What to do

- **Thematic groupings**, but not semantic ones
- Have learners **generate example sentences** meaningful to them.
- **Form variability** leads to better learning (different speakers, **IPA**, transliteration in the L1 script).
- **Meaning variability** (e.g., multiple meanings or even multiple pictures of an object) at initial learning leads to poorer learning.
- For initial form-meaning learning, encourage **mindful translation** with, yes, a (good-quality) **bilingual dictionary!**
- Once learners have matched a form and a meaning, most learning will happen through **massive exposure to language** and **massive practice!** Vary challenging and easy texts (the latter for **fluency** practice).



WHAT RESEARCH CAN TEACH US ABOUT TEACHING SECOND LANGUAGE VOCABULARY

Vedran Dronjic
vedran.dronjic@nau.edu

Should students use monolingual or bilingual dictionaries?

- Laufer & Melamed (1994): 3 dictionaries – monolingual, bilingual, and bilingualized
- Unskilled dictionary users (“pre-advanced” level):
 - **COMPREHENSION: bilingualized > monolingual**
 - **PRODUCTION: bilingual, bilingualized > monolingual**
- Average users:
 - **COMPREHENSION: bilingualized > bilingual**
 - **PRODUCTION: bilingualized > monolingual**
- Good users:
 - **fine with any of the three dictionaries (trend toward bilingualized > monolingual > bilingual)**