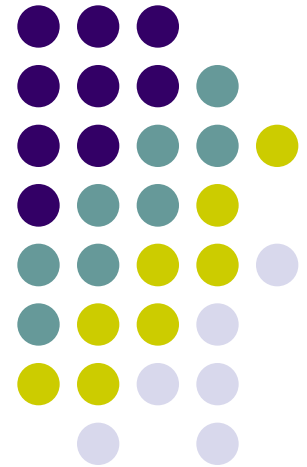


Towards a genre-based approach to the lexicographical treatment of phraseology in electronic monolingual learners' dictionaries

Magali Paquot

Centre for English Corpus Linguistics

Université catholique de Louvain





Background

- Monolingual learners' dictionaries (MLDs) have taken “more proactive steps to help learners negotiate known areas of difficulty” (Rundell, 1999: 47), to the point that they are now conceived as comprehensive writing tools.
- Possibility of including more information on collocations and other phraseological units.



Objective

- Investigate the usefulness of phraseological information in electronic monolingual learners' dictionaries for academic writing
 - “a shared scientific voice or ‘phraseological accent’” (Gledhill, 2000: 204)
 - [= QUALITY]



Focus on lexical verbs

- “discuss matters lying at the very heart of the scholarly process” (Meyer, 1997: 368)
- create a minefield of difficulties for non-native writers (e.g. Hinkel, 2002)
- Each lexical verb has its own preferred lexico-grammatical company in academic discourse (cf. Granger & Paquot, 2008)
 - subjects (*this study shows that, the evidence suggests that; these results suggest that*),
 - objects (*support the view / hypothesis that ..., provide evidence / information*)
 - adverbs (*differ significantly; vary considerably / widely; apply equally; closely related; widely used; generally accepted*)
 - routinized structures (*as discussed in; there is (no, some, little) evidence that, it should be noted that*).



Verbs of evidence

- *argue, attest, confirm, contradict, corroborate, demonstrate, disprove, evidence, evince, illustrate, imply, indicate, mean, point to, prove, reveal, show, substantiate, suggest, support, tell, testify, verify*
- Academic Keyword List (Paquot, 2010)
 - 930 potential academic words, i.e. words that are reasonably frequent in a wide range of academic texts but relatively uncommon in other kinds of texts and which, as such, might be used to refer to those activities that characterize academic work, organize scientific discourse and build the rhetoric of academic texts, and so be granted the status of academic vocabulary.

Collocation boxes in the 'Big Five'



- [CALD] Cambridge Advanced Learner's Dictionary (3rd edition), 2008.
- [OALD] Oxford Advanced Learner's Dictionary (8th edition) 2010.
- ***[CCAD] Collins-Cobuild Advanced Learner's English Dictionary (6th edition), 2009.***
- ***[MEDAL] Macmillan English dictionary for advanced learners (2nd edition), 2007.***
- [LDOCE] Longman Dictionary of Contemporary English (5th edition), 2009.



Methodology

- Collocation boxes
 - COBUILD / MEDAL / LDOCE
- Coverage of the most typical academic collocations
- Word Sketches in the *Corpus of Academic Journal Articles (CAJA)* (Kosem, 2010)
 - 10 most frequent collocates per category
 - V + N / N + V structures

Word Sketch



Sketch Engine Home Settings Change password Log out

Search in Help

user: Dr. Magali Paquot corpus: Corpus of Academic Journal Articles (CAJA) Search in Corpus of Academic Journal A

Concordance

Word List

Word Sketch

Thesaurus

Find X

Sketch-Diff

? Help on main menu

Save

Change options

Turn on

clustering

More data

Less data

Switch menu position

argue (verb)

Corpus of Academic Journal Articles (CAJA) freq = 27628 (292.8 per million)

PP_PP_between	44	2.5	PP_against-i	405	41.7	subj_NP	7039	3.7	PP_for-i	1127	4.8	AVP_mod	2210	4.2
-i			view	20	19.37	article	201	33.8	importance	31	22.1	also	437	38.4
for	39	28.68	possibility	18	20.97	paper	165	29.47	existence	24	21.75	as	190	27.93
			interpretation	13	17.23	other	119	27.1	view	22	15.75	further	130	37.89
			idea	12	15.25	author	113	29.11	role	19	12.41	not	128	16.46
			hypothesis	10	14.52	researcher	94	27.45	approach	18	11.5	then	80	24.41
			mechanism	9	12.16	scholar	89	30.08	claim	16	15.17	recently	78	33.36
			role	9	10.46	Kant	56	30.19	need	16	14.45	often	68	24.13
			use	9	10.35	critic	50	27.04	presence	15	13.72	convincingly	64	49.8
			assumption	8	12.76	DOJ	47	40.13	necessity	12	19.9	even	56	20.83
			analysis	8	7.49	Schmalensee	44	39.89	interpretation	12	12.66	now	55	24.21
			effect	8	6.86	theory	44	10.75	model	12	5.15	therefore	42	21.49
			notion	7	12.41	study	40	4.32	use	11	8.21	already	38	21.85
			model	6	4.84	theorist	35	22.85	change	11	6.81	strongly	36	23.51
			explanation	5	9.86	literature	35	12.78	conception	10	14.59	thus	28	16.06
			position	5	7.43	government	33	11.22	theory	9	7.32	similarly	27	24.21

Dictionary look-up as information retrieval



- Retrieval [= search] effectiveness
 - **Recall**: proportion of relevant materials retrieved
 - **Precision**: proportion of retrieved materials that are relevant

Salton (1989); Ponte & Croft (1998); Granger (1997); Shaffi & Rather (2005)

Recall



$$\text{Recall rate of 'best' academic collocations} = \frac{\text{Number of best academic collocations}}{\text{Best academic collocations}}$$

- Quantifying how well a dictionary answers specific users' needs
- DEMONSTRATE: *effect, **importance**, difference, relationship, **ability**, role, level, increase, effectiveness, potential (20%)*
- 0% (N + 'prove', LDOCE) → 90% (N + 'show', LDOCE)
- Mean: 32.5%



Recall in MEDAL

- ‘Collocations’ boxes for specific senses
 - *argue, show, suggest, support*
- Relational collocations (e.g. ‘as object’) (Evert, 2004)

- 3c. if land supports people or animals, enough food grows on it to feed them
- 4 to show that an idea, statement, theory etc is true or correct
Several witnesses supported Mrs Carson’s claim of harassment.
Our conclusions are supported by extensive research.
You will need to provide supporting evidence for your claim.

Collocations

Extra information

Collocations

Nouns frequently used as objects of *support* 4

- argument, claim, conclusion, contention, hypothesis, idea, theory, view

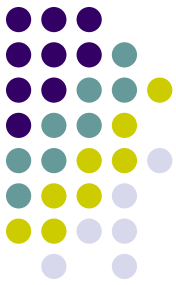
Precision



Precision rate of ‘best’ academic collocations =
$$\frac{\text{Number of best academic collocations}}{\text{Number of collocations}}$$

- ‘support’ (19.4%, LDOCE)
 - ***claim**, development, effort, **evidence**, family, government, **hypothesis**, **idea**, party, policy, project, proposal, etc*
- 6.8% (‘prove’, LDOCE) → 75% (‘support’, MEDAL)
- Mean LDOCE: 23.9%

LDOCE



LDOCE 5 Home | Copy | Print | My Dictionary | Help

Longman Dictionary of Contemporary English

Dictionary | Activator | Grammar | Exercises | Vocabulary Trainer | Teacher Resources | Pop up Dictionary | Writing Assistant

suggest
Spell check
Phrase search
Pronunciation search
Advanced search

A-Z **CULTURE**

suggest
suggestible
suggestion
suggestive
suggestively

SEARCH RESULTS
suggest verb

sug-gest **S1 W1** /sə'dʒest \$ səg'dʒest/ verb [transitive]
[Word family](#) [Word origin](#) [Verb form](#)

1 to tell someone your ideas about what they should do, where they should go etc → **propose**:
• The zoo asked its visitors to suggest a name for the new baby panda.
suggest (that)
• Her mother suggested that she should go and see the doctor.
• I suggest you call him first.
• **It has been suggested that** the manager will resign if any more players are sold.
suggest doing something
• Tracey suggested meeting for a drink after work.
suggest how/where/what etc
• Can you suggest what kind of tool I should use?
can/may I suggest (=used to politely suggest a different idea)

COLLOCATIONS
Collocations from this entry
Collocations from other entries
Collocations from the corpus

THESAURUS
Thesaurus
Longman Language Activator
Word sets

PHRASE BANK
Phrases from this entry
Phrases from other entries

EXAMPLE BANK
Other dictionary examples
Examples from the corpus

PRONUNCIATION
Practice your pronunciation

USER NOTES
Add a note
View note

Three collocation boxes for 'suggest'



COLLOCATIONS FROM THIS ENTRY	COLLOCATIONS FROM OTHER ENTRIES	COLLOCATIONS FROM THE CORPUS
<ul style="list-style-type: none">→ can/may I suggest→ evidence/results/data/studies etc suggest(s) that→ It has been suggested that	<ul style="list-style-type: none">→ a poll shows/indicates/suggests sth→ a study suggests/indicates sth→ a survey suggests sth→ an analysis suggests/indicates sth→ common sense suggests sth→ data indicates/suggests sth→ experience suggests/shows sth→ logic suggests sth→ statistics suggest sth→ sth suggests a connection→ strongly suggest/advise/recommend sth→ suggest an alternative→ suggest/put forward a solution	<p>ADVERB</p> <ul style="list-style-type: none">→ also→ strongly <p>NOUN</p> <ul style="list-style-type: none">→ data→ evidence→ experience→ report→ research→ result→ study <p>VERB</p> <ul style="list-style-type: none">→ seem

Three collocation boxes for 'suggest'



COLLOCATIONS FROM THIS ENTRY

- can/may I suggest
- evidence, results/data/studies etc suggest(s) that
- It has been suggested that

COLLOCATIONS FROM OTHER ENTRIES

- a poll shows/indicates/suggests sth
- a study suggests/indicates sth
- a survey suggests sth
- an analysis suggests/indicates sth
- common sense suggests sth
- data indicates/suggests sth
- experience suggests/shows sth
- logic suggests sth
- statistics suggest sth
- sth suggests a connection
- strongly suggest/advise/recommend sth
- suggest an alternative
- suggest/put forward a solution

COLLOCATIONS FROM THE CORPUS

ADVERB

- also
- strongly

NOUN

- data
- evidence
- experience
- report
- research
- result
- study

VERB

- seem

Collocations from other entries: 'prove'



LDOCE5

COLLOCATIONS FROM OTHER ENTRIES

- an experiment shows/proves/demonstrates sth
- be proved right
- be proved wrong
- be/become/prove an attraction
- prove a theory
- prove a will
- prove an embarrassment
- prove an obstacle
- prove correct
- prove costly
- prove crucial
- prove expensive
- prove impossible
- prove irresistible
- prove (to be) unfounded
- prove unsuccessful
- prove useful
- prove useless
- prove valuable
- prove your/a point
- prove/confirm/establish the existence of sth
- proved disastrous

Precision as a measure of information stress or overload



- “A major problem in the present information age is not the absence of data from which the needed information can be retrieved, but the abundance of unstructured data” (Tarp, 2009: 26)



Collocations in MLDs

- PROVE: *prove sb's guilt/innocence, prove one's manhood/masculinity, prove your loyalty, prove your mettle*
- REVEAL: *reveal sb's identity, reveal a secret, reveal the truth*
- SHOW: *show compassion, contempt, desire, courage, emotions, enthusiasm, show a film*
- SUPPORT: *support an accusation, a cause, a charity, an event*

“Inappropriate” collocations for academic writing



- Adverbs!

2 [intransitive/transitive] to give reasons why you believe that something is right or true

Successful economies, she argues, are those with the lowest taxes.

argue for/against: Woolf's report argued for (=supported) an improvement in prison conditions.

argue that: Reuben opposed the new road, arguing that it wasn't worth spending \$25 million to cut seven minutes off drivers' journey times.

Several people stood up to argue against (=say they do not support) moving the students to the new school.

Collocations

A screenshot of a Macmillan English Dictionary window. The title bar reads "Macmillan English Dictionary". Below the title bar is a red header with the text "Extra information". Underneath that is a dark blue header with the text "Collocations". The main content area lists adverbs frequently used with "argue 2":

- consistently, convincingly, forcefully, passionately, persuasively, plausibly, strongly

The limits of reference corpora for specialised lexicography



- OALD: The British National Corpus (c. 100m)
- LDOCE: Longman Corpus Network (c. 330m)
- CCAD: The Bank of English (c. 650m)
 - Different text types and genres

- Selection of collocations
- Sense ordering
- Example sentences (cf. Paquot, 2011)

Example bank for 'prove' in OALD



- He tried to **prove** his theory to his friends.
- I certainly don't have anything to **prove**—my record speaks for itself.
- I was determined to **prove** my critics wrong.
- The deaths are being treated as suspicious until we can **prove** otherwise.
- Their behaviour just **proves** my point.
- This theory cannot be **proved** scientifically.
- What are you trying to **prove**?
- 'I know you're lying.' '**Prove** it!'
- Are you just doing this to **prove** a point?
- Can it be **proved** that he did commit these offences?
- He felt he needed to **prove** his point.
- Just give me a chance and I'll **prove** it to you.

Louvain English for Academic Purposes Dictionary (LEAD)



- Specific population: students and researchers who need to write academic texts in English
- Specific set of vocabulary
 - Academic vocabulary: Academic Keyword List (Paquot, 2010)
 - Nouns: *cause, contrast, hypothesis*
 - Verbs: *argue, suggest, discuss*
 - Adverbs: *however, on the contrary, albeit*
 - Adjectives: *comparable, parallel*
 - Prepositions: *in spite of, unlike, with regard to*

Selection of general academic collocations and lexical bundles



- General EAP corpus
 - Academic component of the BNC (15 million words)

C E C L The Louvain EAP dictionary

Search Selected discipline: *Business*; selected mother tongue: *French* ([change settings](#))

[Home](#) [Search by function](#) [Search by translation](#) [Corpus search](#) [Exercises](#)

Headwords illustrate (vb.)	illustrate (vb.)	Collocations
Function	to show the meaning of something more clearly, often through the use of an example:	<i>illustrate</i> + N <i>point, importance, problem, principle, way, difficulty, difference, method, type, use, process, nature, approach, effect, situation, pattern, range, feature, aspect, complexity, procedure, idea, issue</i>
Give examples	<i>The 2004 tsunami in Southeast Asia illustrates the need for collaborative regional hazard surveillance.</i>	<i>illustrate</i> + Adv <i>well, also, best, clearly, here, better, below, further, nicely, only, vividly, above, neatly, just, simply, thus, perfectly, strikingly</i>
In French	In academic writing, the verb illustrate is mostly used in the active structure with subjects such as <i>case, example, figure</i> , and <i>study</i> , or in the passive:	Lexical bundles
illustrer montrer donner un exemple de être un exemple de	<ul style="list-style-type: none">• <i>This case illustrates the sense of hopelessness that may develop in the unemployed, especially when finding a job appears to be an almost impossible task.</i>• <i>This research programme illustrates the part that accident can play in science.</i>• <i>The interplay of these factors is illustrated by the following case studies.</i>• <i>As illustrated in the first chapter, formal community care services are provided by a variety of different agencies.</i>• <i>Third party interests in a dispute will range in intensity and directness, as illustrated by the background to Nicaragua v. the United States.</i>	<i>illustrate how</i> <i>illustrate that</i> <i>this can be illustrated by</i> <i>these examples illustrate</i> <i>this is well illustrated by</i> ...



Customisability in LEAD

- Discipline
- Mother-tongue background

C E C L

The Louvain EAP dictionary

Please select a discipline:

What is your mother tongue?

Contact : [Magali Paquot](#) & [Sylviane Granger](#)®

EAP dictionary
[Welcome](#)
[Concordancer](#)
[Exercises](#)
[References](#)

Links
[Centre for English Corpus Linguistics](#)

Intranet
[Lexicographer's corner](#)

Automatic selection of discipline-specific examples



- Sentence examples automatically extracted from discipline-specific corpora (Paquot, forthcoming)
 - Business, medicine, linguistics

*In their entirety, our results **illustrate** the **importance** of time and temporal variables for understanding both the behaviors of entrepreneurs and the contexts in which entrepreneurs work. (BUS)*

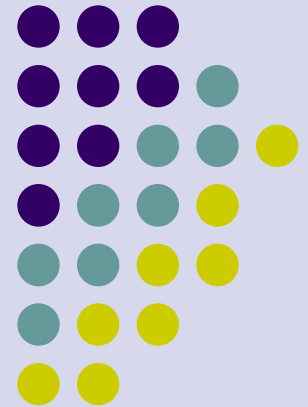
*These findings between temporal variables and other important human outcomes (see Blueborn, 2002 for discussions of many such relationships) not only **illustrate** the **importance** of studying time to understand human affairs, but they suggest time is likely to be an important factor for entrepreneurs and their enterprises. (BUS)*

*Taken as a whole, our results **illustrate** the **importance** of time and temporal variables for understanding the behaviors of entrepreneurs and the contexts in which entrepreneurs work. (BUS)*

*In order to **illustrate** the **importance** of the tacit knowledge component in supply chain management and outsourcing, let us look at a hypothetical example of an automaker designing and implementing an outsourcing strategy for its cooling system. (BUS)*

*While it has not been examined in relation to trust in leadership, Edmondson (2003) found that team leaders could develop psychological safety within the team through interpersonal activities which serve to motivate the team and **illustrate** the **importance** of all members' inputs and downplay power differences. (BUS)*

Concluding remarks



Treatment of collocations in electronic learners' dictionaries



- “at the present stage the access options provided for phraseological units in the electronic versions have still not found an adequate balance between the parameters of quantity and quality.” (Götz-Votteler & Herbst, 2009: 57)
- Much remains to be done!



Desiderata

- Systematize collocation boxes
 - Not only for nouns and highly frequent words
- Cover other lexico-grammatical structures
 - *argue + for + importance / existence / view*
 - *indicate + in + table / figure / parenthesis*
 - *suggest + Ving: using / adding / considering / increasing*
 - *suggest + as + cause / explanation / factor*
 - *support + and_or + oppose / encourage / refute / promote*
- Add lexical bundles



Be context sensitive

- Sense-differentiated collocations
- “ Collocational normality is dependent on **genre, register** and **style** i.e. what is normal in one kind of text may be quite unusual in another.” (Partington, 1998: 17)
- “Particularly crucial is the function of phraseological information in relation to the **needs and interests of the target users**”. (Moon 2008: 333)



Pedagogical lexicography

- The monolingual learner dictionary as a “one-size-fits-all package” (Rundell 2007: 50) is probably dead, or at least seriously ill 😞
- Needs to re-invent itself to compete against all the new players, especially online (Lannoy, 2010)
- Customisation
 - Genre, discipline, L1 background

Specialized corpora, customisation and users' needs



- Function theory of lexicography
 - Users' needs
 - Relevance
 - Practical lexicography
 - Specialized corpora
 - Corpus-based data (frequency, preferred environment, etc)
- Concrete extra-lexicographical situations
-

**CORPUS-BASED APPROACH TO
USERS' NEEDS**

Thank you!

magali.paquot@uclouvain.be

