

SkELL: Web Interface for English Language Learning

Vít Baisa and Vít Suchomel

Natural Language Processing Centre
Faculty of Informatics
Masaryk University
and
Lexical Computing Ltd.

5 December 2014



SkELL = Sketch Engine for Language Learning

- A new web interface,
- optimised for mobile and touch devices,
- aimed at teachers and students of English language,
- based on a specially processed corpus data suitable for the language learning purpose,
- spam free, high quality texts from various domains including diverse text types covering majority of English language phenomena.

40 good dictionary example sentences for a given word or phrase.

language learning 1.0 hits per million

- 1 The highly engaging courses utilize progressive **language learning** methods.
- 2 These external characteristics may impact **language learning** opportunities.
- 3 Their **language learning** ability is very strong .
- 4 They are used very widely in **language learning** .
- 5 What is the best **language learning** software?
- 6 The development of **language learning** is thereby disrupted.
- 7 **Language learning** normally occurs most intensively during human childhood .
- 8 The same is true with children with **language learning** difficulties.
- 9 A broader approach to **language learning** than community language learning.
- 10 The importance of listening in **language learning** is often overlooked.

Functionality – Word sketches

Collocational and grammatical profiles for a given word.

lunch *noun* or *verb*

verbs with lunch as object	verbs with lunch as subject	modifiers of lunch	nouns modified by lunch	words and/or lunch
eat pack box cook enjoy cater skip grab serve prepare finish bring	consist break follow serve include adjectives with lunch ready special available	picnic packed breakfast buffet reduced-price leisurely delicious three-course sack sumptuous Sunday nutritious	counter break menu buffet dinner pail basket special snack packing box hour	breakfast dinner snack refreshment brunch supper recess tea coffee drink break meal

Functionality – Word sketches

Interface for mobile devices

vegetable *noun*

verbs with vegetable as object

cook

steam

chop

roast

ferment

eat

blanch

grow

plant

boil

slice

harvest

cultivate

dry

consume

verbs with vegetable as subject

grow

contain

help

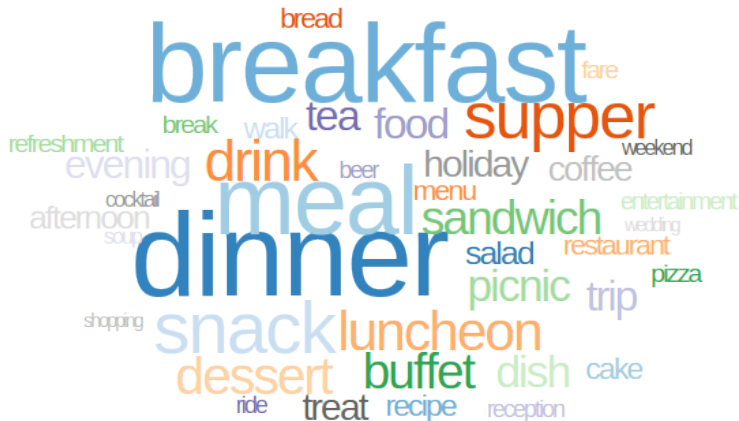
include

be

provide

Functionality – Thesaurus

Up to 40 words similar to (though not necessarily synonymous with) the given word



Sources used in the SkELL corpus

Subcorpus	tokens	used
Wiki	1.6 G	500 M
Gutenberg	530 M	200 M
White	1.6 G	500 M
WebBootCated	105 M	all
BNC	112 M	all
Feed Corpus	344 M	200 M

Processing the data

- ① normalization of quotes and interpunction,
- ② tokenization by `unitok.py`,
- ③ POS tagging by `TreeTagger`,
- ④ deduplication on sentence level by `onion`,
- ⑤ sentences sorted by GDEX 3.

Conclusion

- SkELL = Sketch Engine for Language Learning
- Available for free at <https://skell.sketchengine.co.uk>.