

WordGap – Automatic Generation of Gap-Filling Vocabulary Exercises for Mobile Learning

Susanne Knoop
Sabrina Wilske

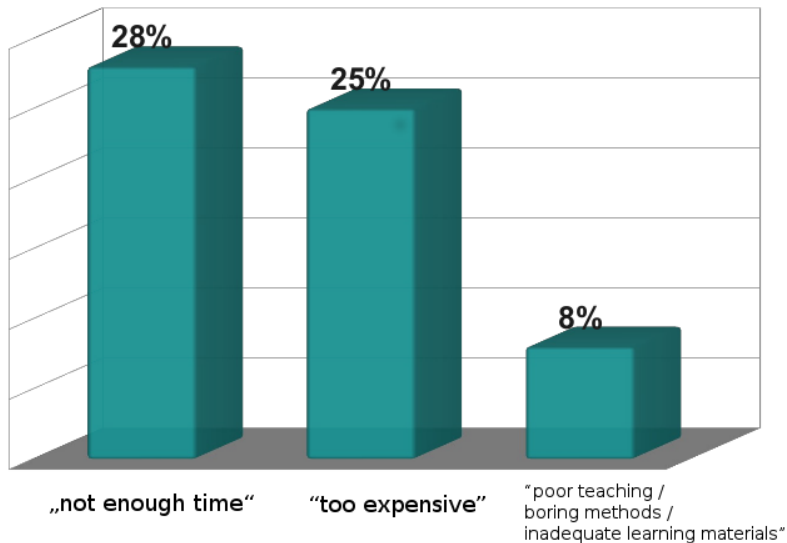
University of Bremen, Germany
Digital Media in Education (Prof. Heidi Schelhowe)

May 22, 2013

Overview

- 1 Motivation
- 2 Previous Work
- 3 WordGap
- 4 NLP Tools
- 5 Outlook

Obstacles To Language Learning



European Union, Special Eurobarometer 386

The Solution

- ① Free CALL software
- ② Mobile apps (“learn everywhere, everytime”)
- ③ Exercises from texts that are of special interest to the user

Vocabulary Learning With Multiple-Choice Cloze Exercises

- Cloze exercises consist of:
 - ① A sentence
 - ② A target word that has been replaced by a gap
 - ③ (normally three) distractors to distract from the target response
- They are widely used in second language training and testing

The Challenge

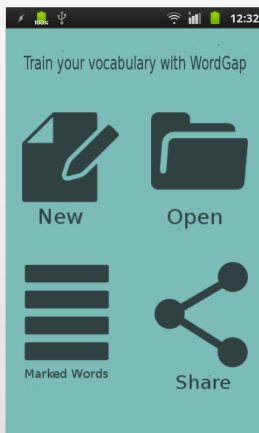
- It normally requires expert knowledge to create cloze exercises:
 - ① Distractors of multiple-choice cloze exercises should not be totally unlikely but also not too similar to the target word
 - ② Distractors should have the same grammatical form as the target word

Previous Work

- Coniam (1997), Brown et al. (2005): Distractors of the same part of speech and similar frequency as the target word
- Sumita et al. (2005): Distractors are false synonyms of the target word (verified by a Google search)
- Evaluation: Performance of students in the automatically generated tests correlates strongly with standard vocabulary tests.

WordGap

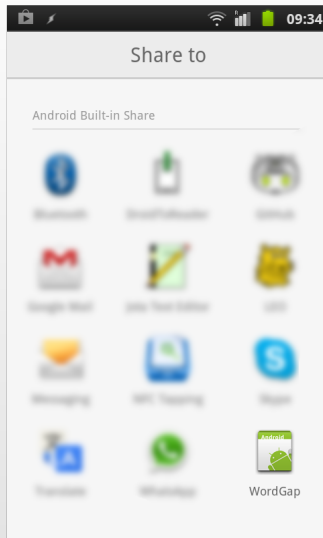
- server component and mobile app for Android smartphones
- informal vocabulary learning, adapted to the learner's interest
- multiple-choice cloze exercises from a text file or webpage



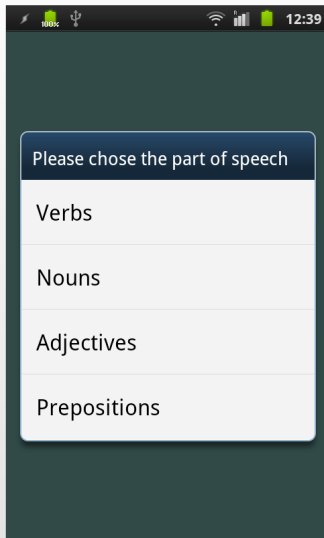
Share a webpage from the browser



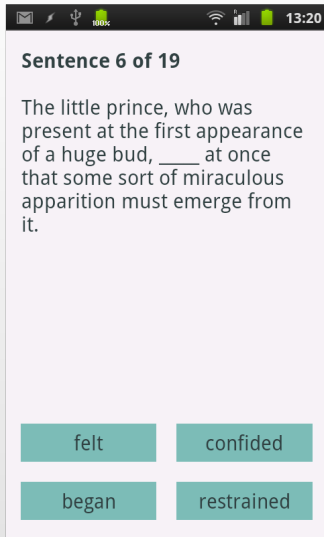
Send the webpage to WordGap



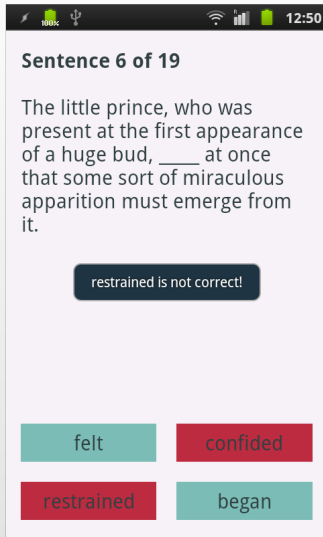
Choose a part of speech



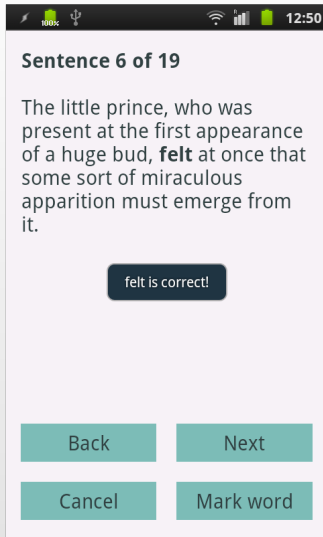
Display the exercise sentence by sentence. . .



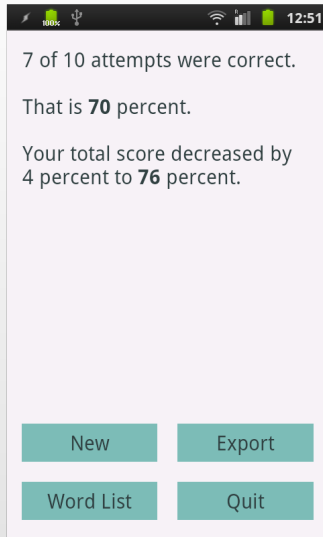
...receive immediate feedback...



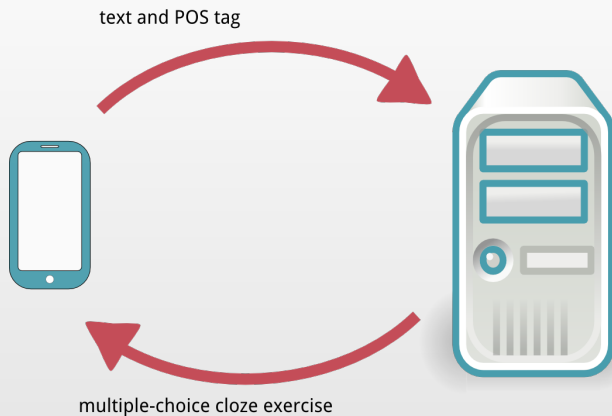
...receive immediate feedback...



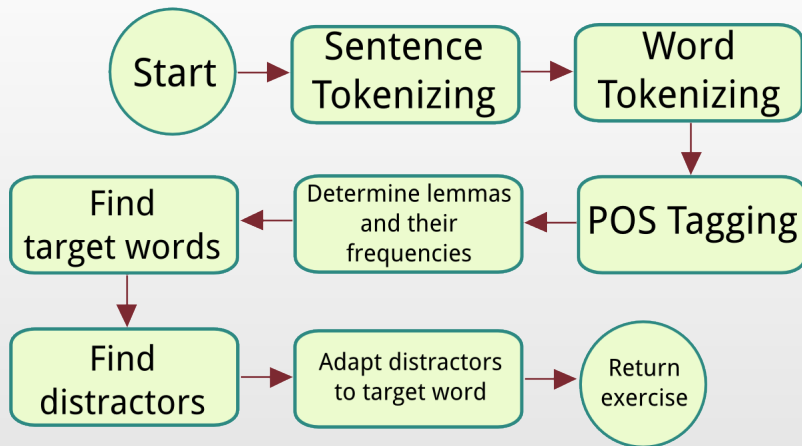
...get the result.



The server-client model of WordGap



The NLP pipeline of the WordGap server



django

NLTK



jsoup



- Hierarchic semantic network of the English language (Miller and Fellbaum, 2007)
- WordGap: find false synonyms or antonyms of a target word as distractors
- False synonyms: synonyms of synonyms or words with the same hypernym
- Distractors from the text's context often more suitable than those from WordNet

Natural Language Tool Kit

- extensive open-source Python library for NLP
- WordGap: sentence and word tokenizing and part of speech tagging
- Naive Bayes Tagger with a cut-off probability of 0.95

- collection of open-source Python libraries for NLP
- WordGap: generate different grammatical forms to adapt distractors to the target word

Outlook

- Include grammatical exercises, e.g. different verb tenses
- Extend to other target languages