Further development of a web application for the annotation of specific linguistic features

Donghwa Park

Informatics
Eberhard Karls Universität Tübingen

Presentation, 08.04.2019

Table of Contents

1. Motivation
2. Approach
3. Evaluation
4. Bonus

Table of Contents

1 Motivation

2 Approach

3 Evaluation

4 Bonus

Motivation

- Many children with reading and writing disorders
- Recognition of syllable important
- Syllable-based reading exercise can help
- Web application considered that generates such exercise texts
- Aimed to help therapists, teachers, parents and children

Starting position

- COAST: Web application that was developed with the previously described motivation
- Problem: It only supports German, some more helpful functions (dictionary, audio pronunciation) can be added and the usability can be improved

Tasks

- English support using CELEX2 (that was used for German)
- Dictionary Service (Definition function & Audio pronunciation function)
- Other modifications to improve usability

Table of Contents

1 Motivation

2 Approach

3 Evaluation

4 Bonus

Dictionary Service

- Definition function
- Audio pronunciation function
- For analyzed text (in the pop-up annotation)
- Separate dictionary service for free look-up

Definition function

- WordNet: Lexical database of English from Princeton University Department of Computer Science
- WordNet as words' definition database
- Database creation using SQLAlchemy
- Python script 'DefinitionService' to get definition and synonyms
- For frontend, 'dictionary' function to send HTTP Request

Audio pronunciation function

- Included in dictionary service
- gTTS (Google Text-to-Speech) python library
- Audio data is binary
- For HTTP Request json format is used
- Json doesn't directly support binary data
- String $\rightarrow$ MP3 file $\rightarrow$ Audio data (binary) $\rightarrow$ String in Base 64

Dictionary Service backend

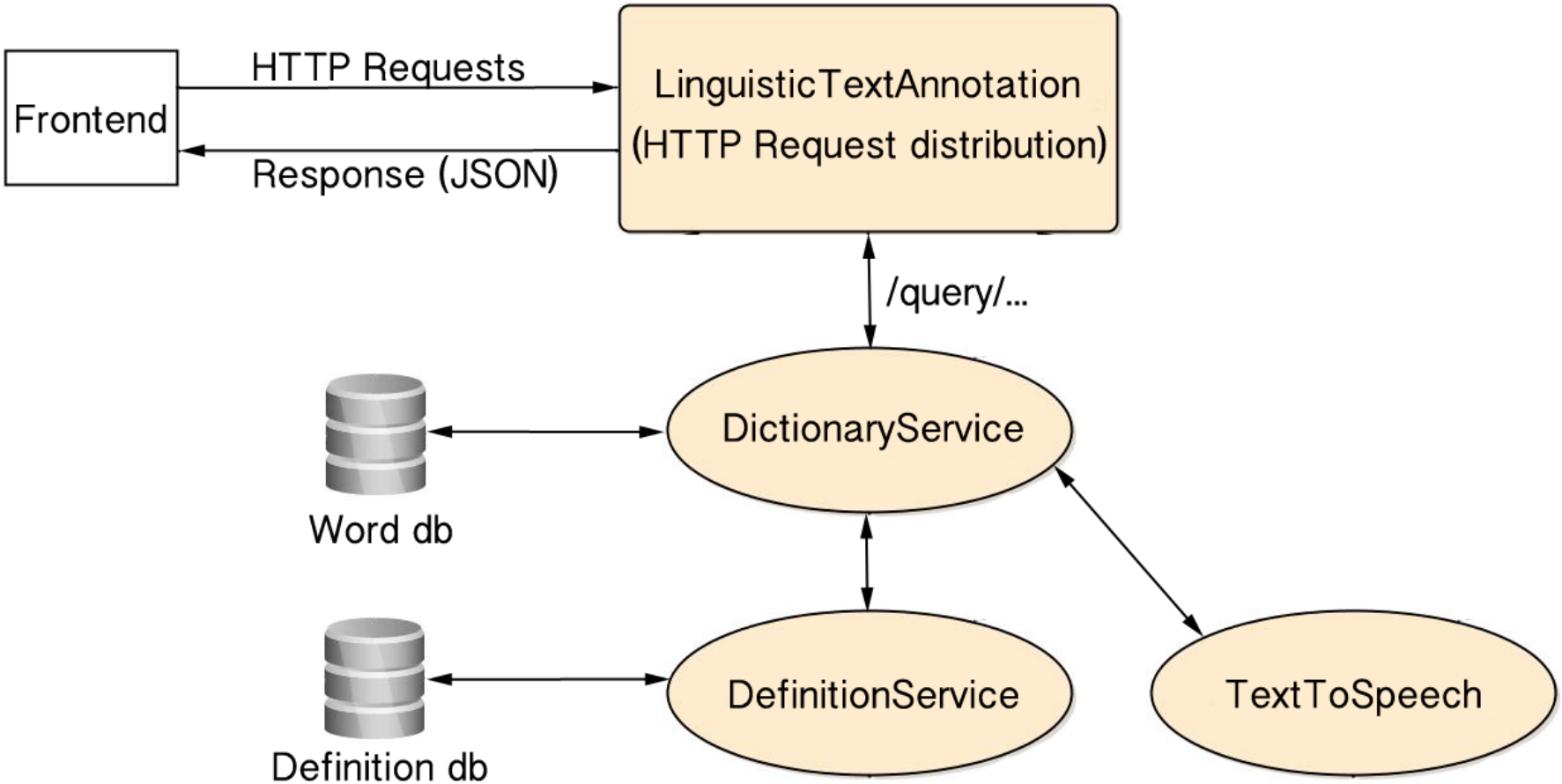


Table of Contents

1 Motivation

2 Approach

3 Evaluation

4 Bonus

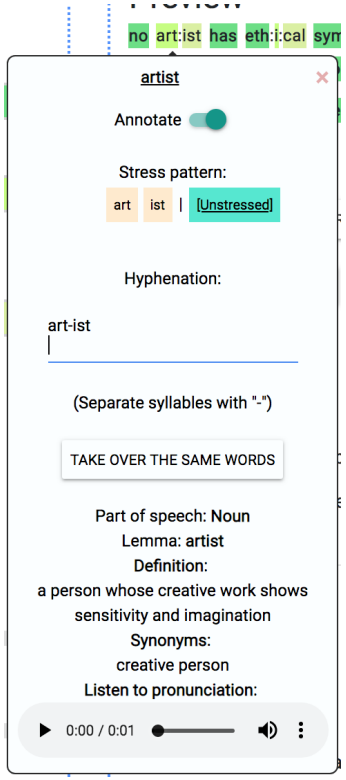
English text generation

No artist has ethical sympathies . An ethical sympathy in an artist is an unpardonable mannerism of style . Thought and language are to the artist instruments of an art . Vice and virtue are to the artist materials for an art .
Diversity of opinion about a work of art shows that the work is new , complex , and vital . When critics disagree , the artist is in accord with himself .

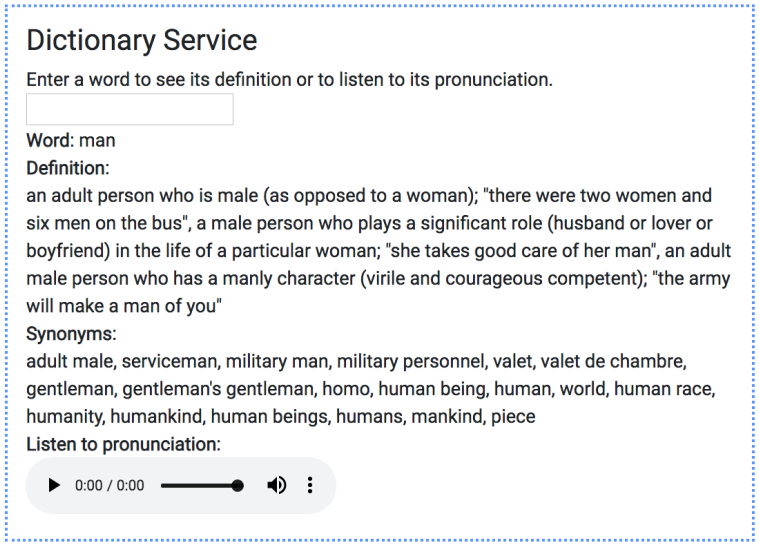
No art-ist has eth-i-cal sym-pa-thies . An eth-i-cal sym-pa-thy in an art-ist is an un-par-don-able man-ner-ism of style . Thought and lan-guage are to the art-ist in-stru-ments of an art . Vice and vir-tue are to the art-ist ma-te-ri-als for an art .
Di-ver-si-ty of o-pin-ion a-bout a work of art shows that the work is new , com-plex , and vi-tal . When crit-ics dis-a-gree , the art-ist is in ac-cord with him-self .

Dictionary service

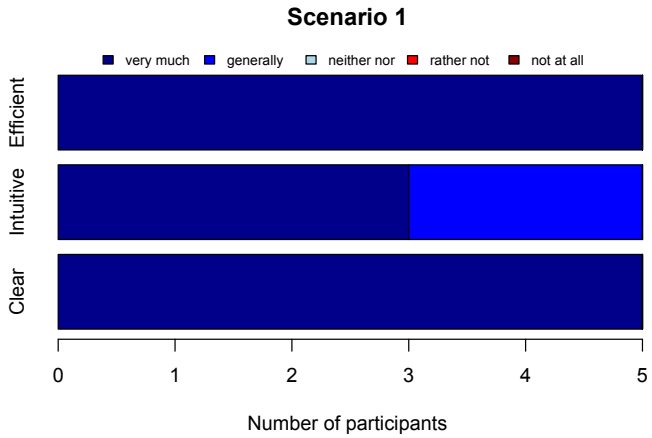
on annotation

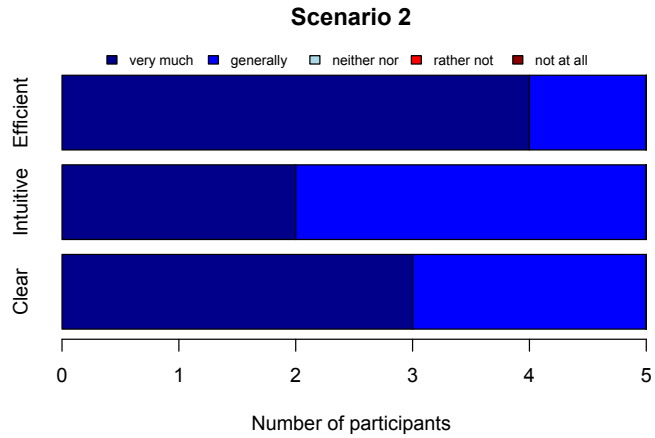


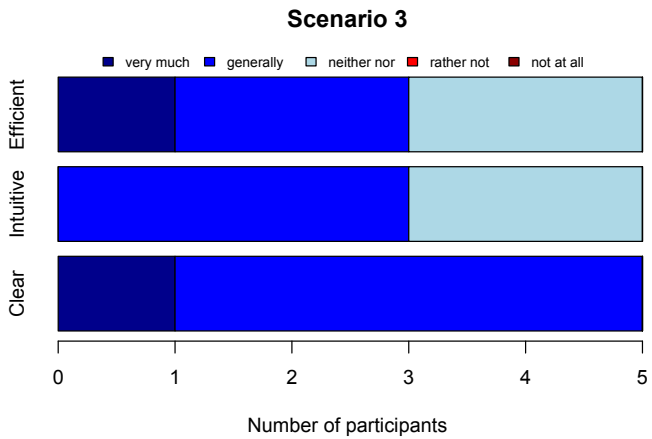
separate service

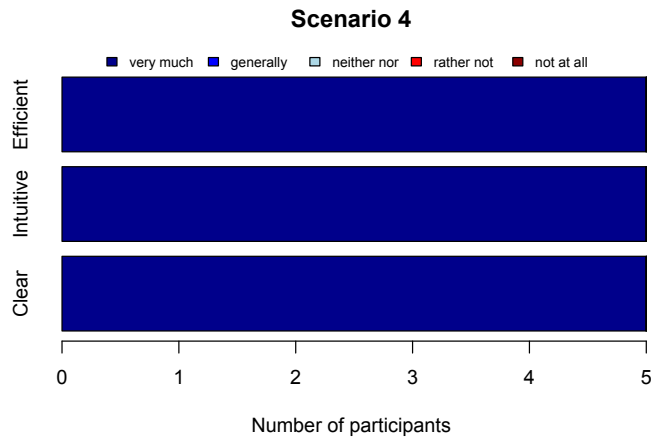


User test


Scenario 1
very much
generally
neither nor
rather not
not at all
Efficient
Intuitive
Clear
0
1
2
3
4
5
Number of participants


Scenario 2
very much
generally
neither nor
rather not
not at all
Efficient
Intuitive
Clear
0
1
2
3
4
5
Number of participants


Scenario 3
very much
generally
neither nor
rather not
not at all
Efficient
Intuitive
Clear
0
1
2
3
4
5
Number of participants


Scenario 4
very much
generally
neither nor
rather not
not at all
Efficient
Intuitive
Clear
0
1
2
3
4
5
Number of participants

Conclusion

- Generation of English exercise texts successful
- Usability improved through modifications
- Font settings can be more approachable
- Further development in other languages possible
- Word database can be extended

Table of Contents

1. Motivation
2. Approach
3. Evaluation
4. Bonus

How to deploy

- Install Python3, Flask, SQLAlchemy, SQLite, Pyphen, Spacy, gTTS on the server
- Deploy the backend
- Adjust variable SERVER_URL in app_service.dart
- Make the frontend deployable using command 'webdev build' and deploy the files under build/web directory (See here and here for more information.)

Thanks for your attention!
Questions & Live-demo