# PV254 - Idioms

Michaela Bamburová, Peter Balčirák,
Štefan Svinčiak, Eva Šmijáková

# Motivation

- Educational system - Teach new idioms
  - Recommend learning idioms of their knowledge level
    - Do not bore a user by teaching them idioms they already know
    - Do not teach them too difficult idioms, they are never going to use

- Target user:

  *"I can speak English already, but I want to improve." (Level B2-C1)*

# Application

- Infinite multiple-choice quiz
  - Question = Idiom
  - Choices = 3 explanations of idiom meaning, choose the correct one
  - After answering, user sees which answer was correct
- Problem of choosing appropriate answers to questions
  - Prevent answering question using elimination method
  - Prevent multiple answers which could be considered correct
- Problem of choosing appropriately hard questions
  - Prevent choosing too hard or too easy questions

# App versions

| Questions | Answers |
|:---:|:---:|
| Random | Random |
| Trained | Random |
| Random | Trained |
| Trained | Trained |

- Trained versions use data gathered by random versions
- Implicit feedback
  - Success rate, number of answered questions, etc. (more in the evaluation part)
- Explicit feedback
  - Very short survey in the end of the quiz

# Implementation

- Backend & data processing:
  - Python
- Frontend:
  - React, Redux, bootstrap
- Infrastructure:
  - PostgreSQL database
  - Deployed on Heroku
  - GitHub repo

# App screenshots

Introduction

What is waiting for you

We have prepared a few questions for you. Each question contains an idiomatic expression together with 3 possible meanings. Your task is to choose the correct one. There is only one correct answer.

Take as much time as you need, but it would help us if you answered at least 10 questions.

When you are done, please click on finish button.

Last but not least

At the end of this quiz, there is a link to a short **survey**. We would be very happy if you fill it.

Many thanks!

Got it

Have kittens

To be very worried, upset or angry about something.

This idiom means to do something bad and not get punished for it.

Said about somone who is honest and good.

Finish

# Data

# Idioms

- Source
  - https://www.myenglishpages.com/site_php_files/vocabulary-lesson-idioms-alphabetical-order.php
  - Web scraping
- Number of idioms
  - 1030
- Structure

```
{
    "category":"nature",
    "definition":"Said about a new and imaginative approach, a change that feels good.",
    "example":"The president says that the country needs a breath of fresh air.",
    "text":"a breath of fresh air"
}
```

# Additional data - corpus

- Source
  - http://www.cs.jhu.edu/~mdredze/datasets/sentiment/ [unprocessed.tar.gz]
- Product reviews from Amazon.com
- Used for counting the frequency of idioms
- Idioms with higher frequency are more common and often easier
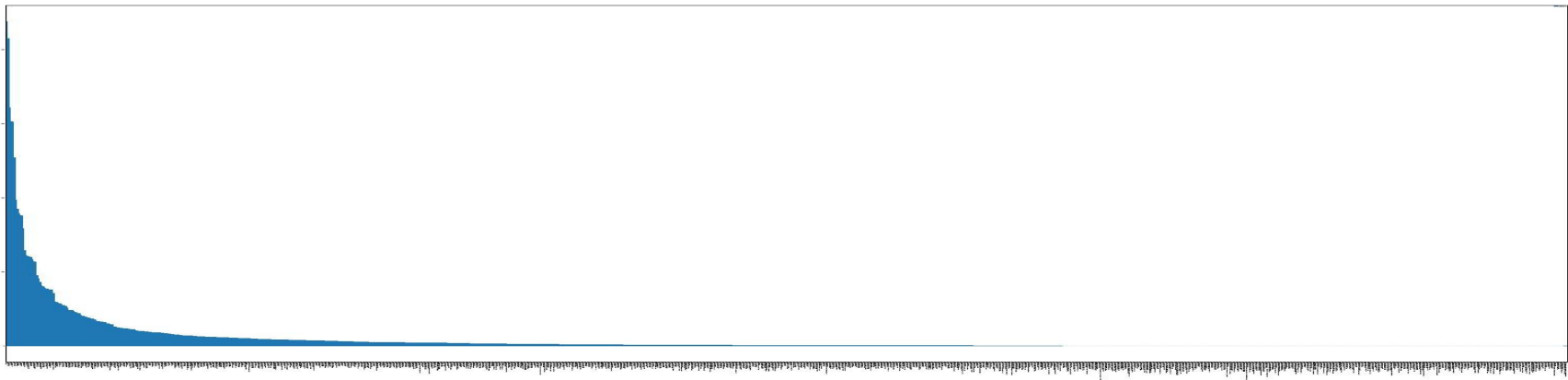
# Models

# Idioms recommendations

- Prevent choosing too hard or too easy questions
- Combine two features:
    - Difficulty of idioms
        - Based on frequency obtained by mentioned corpus
    - User's current success rate
        - Based on intervals with goal to gain and maintain success rate around 80%
        - More correct answers lead to more difficult questions and vice versa

# Answers generator

- Prevent answering by elimination method and multiple correct answers
- Combine three features:
  - Idioms of the same category
  - Previous incorrect answer of other users
  - Cosine distance of explanations based on TF-IDF statistics
    - Discard top 1% similar pairs of idioms to prevent multiple correct answers
    - Example:

| beat the clock | To do something before a deadline |
|---|---|
| against the clock | This idiom means to do something as fast as possible and try to finish it before a deadline. |

# Categories and similarities
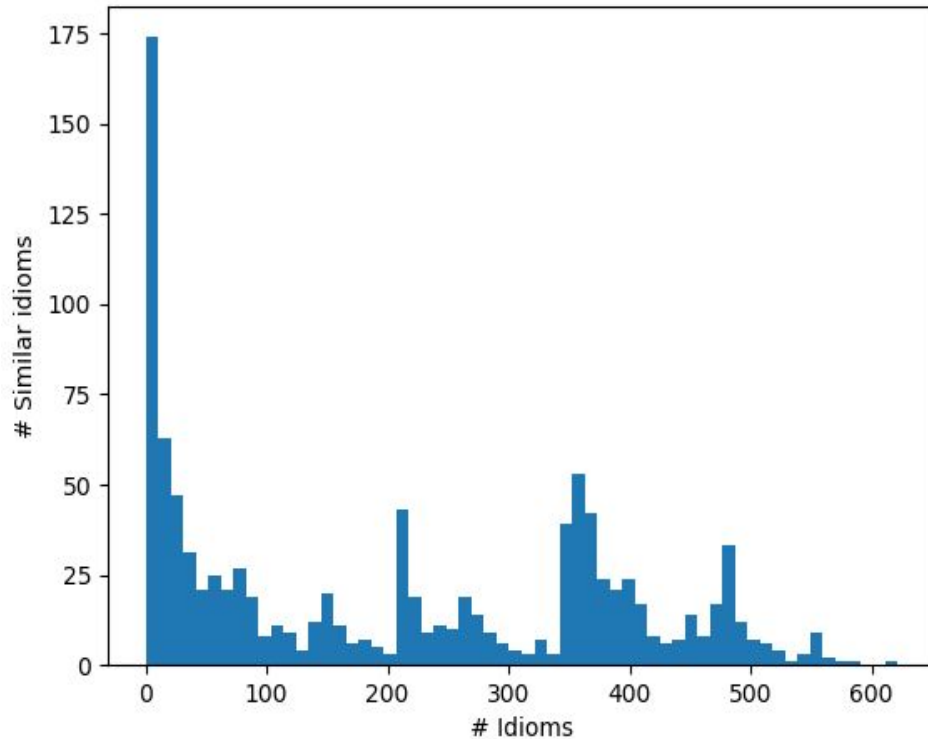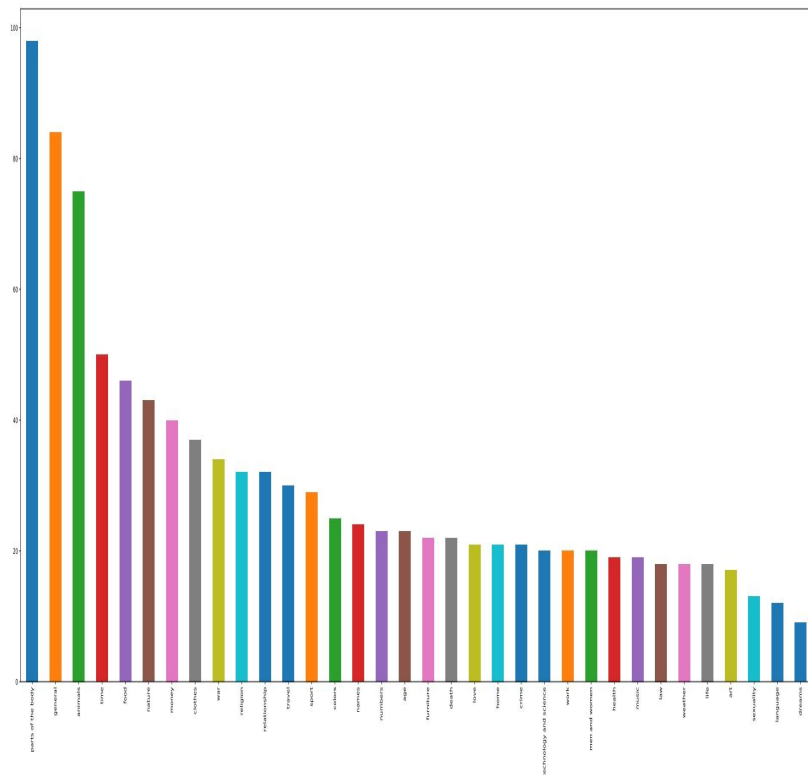
# Results

# Basic statistics

- 239 total users
- 2112 answered questions
- 103 users with more than 5 answered questions
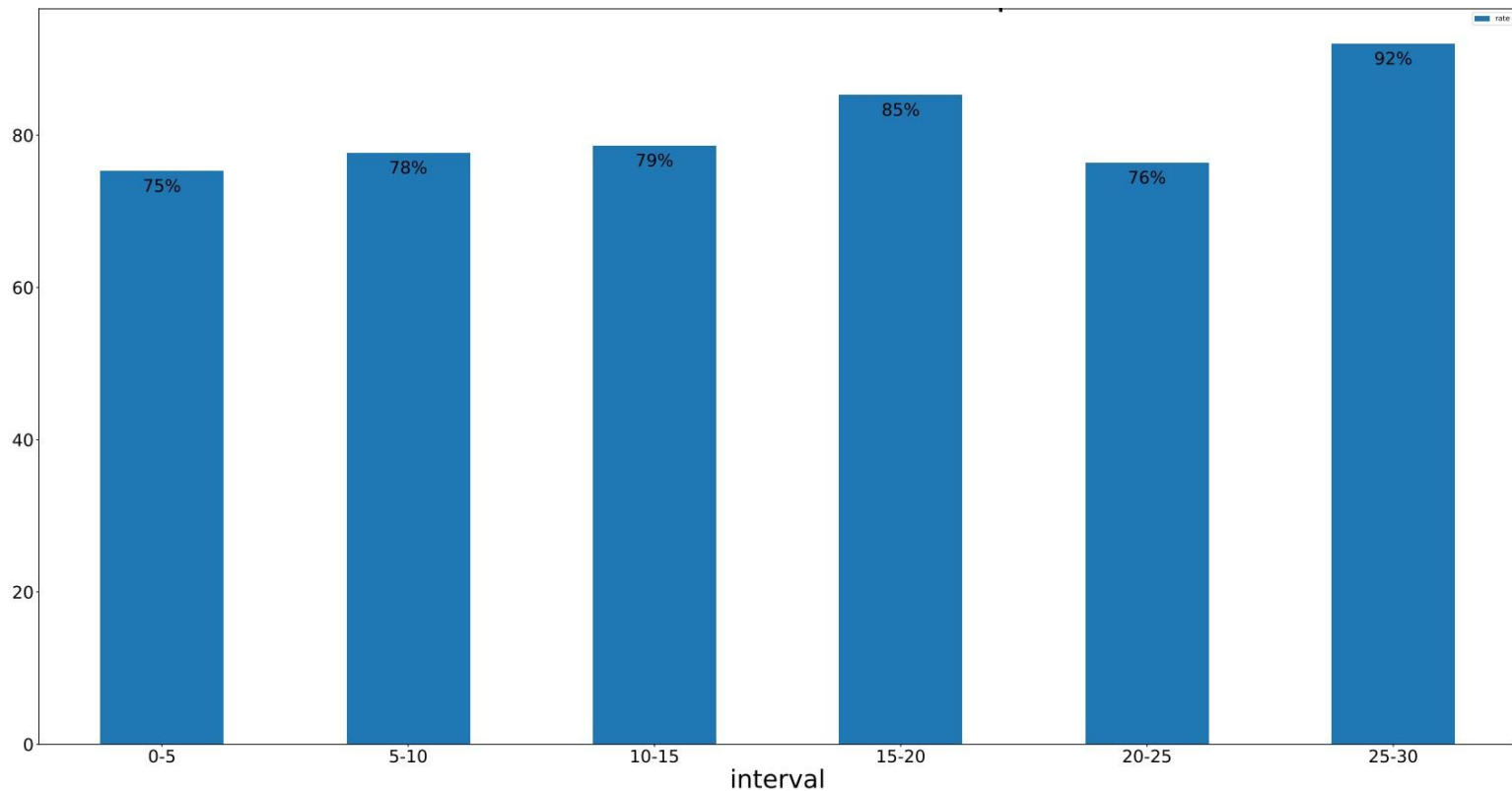- 102 maximum answered questions by a user

# Answered questions by versions
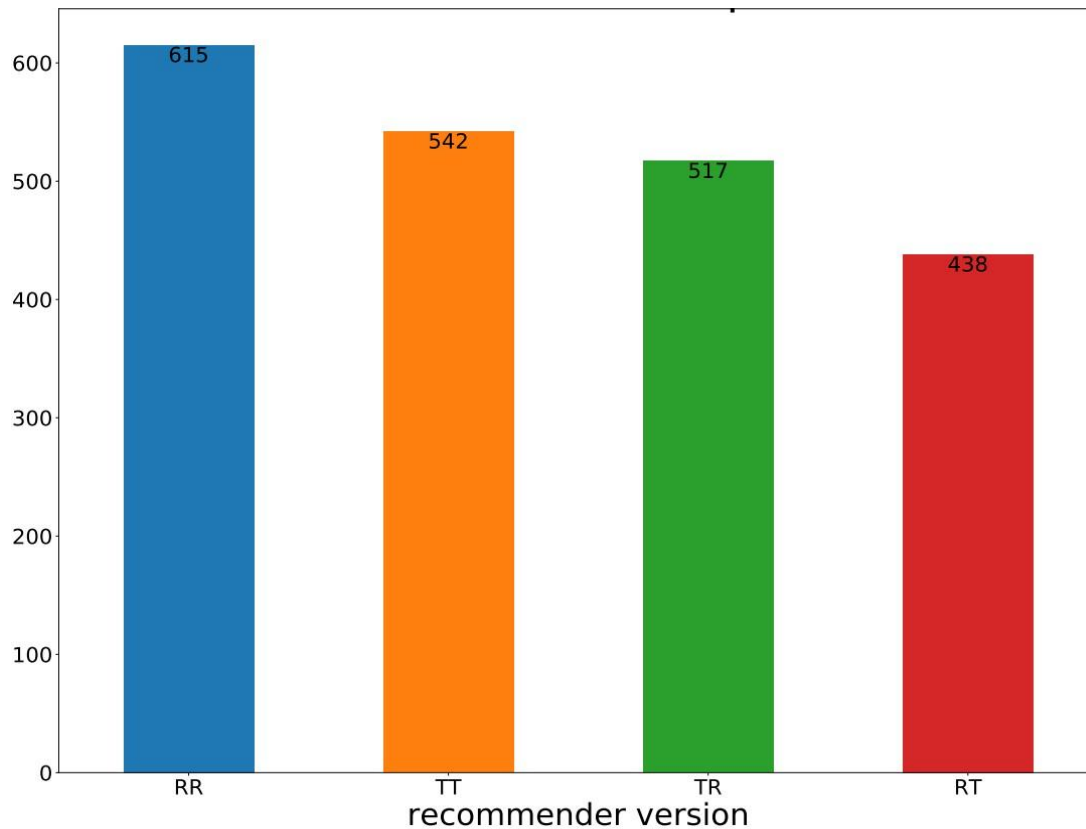
# Success rate
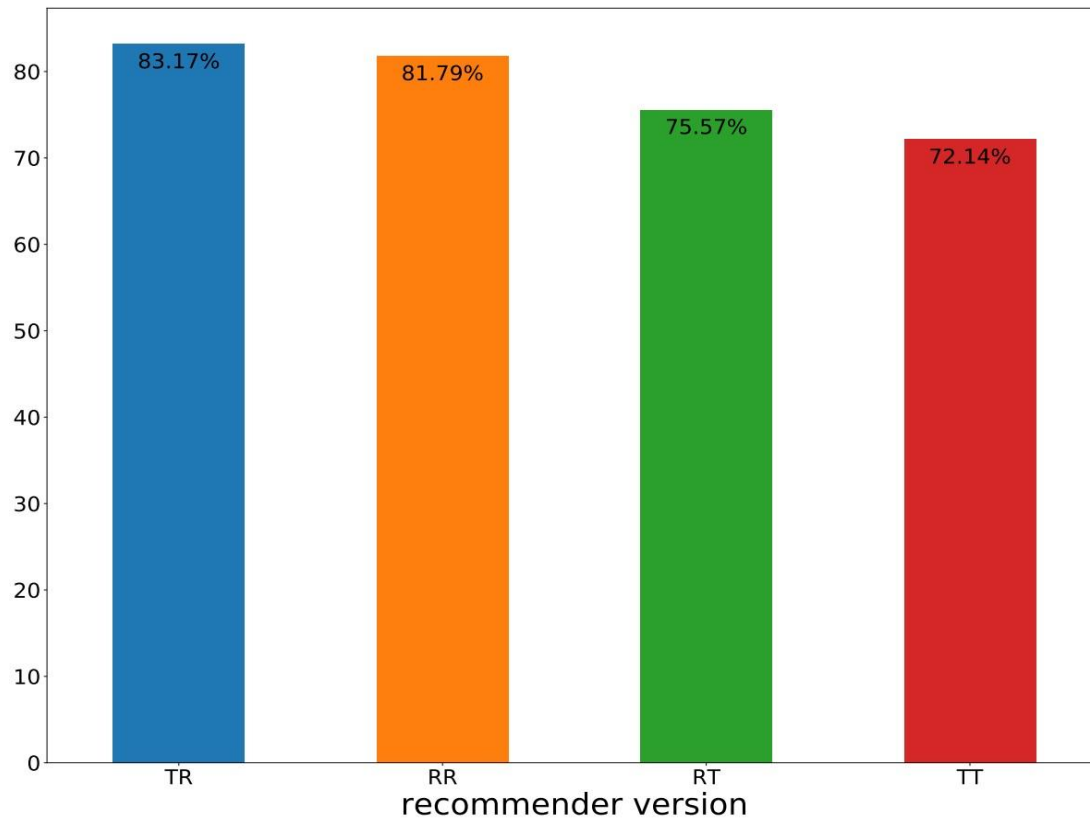
# Success rate - T questions in intervals

# Survey

# Google forms

- 2 questions

## Idiom recommender feedback

* Povinné

**How difficult was the quiz?** *

○ Hard

○ Medium

○ Easy

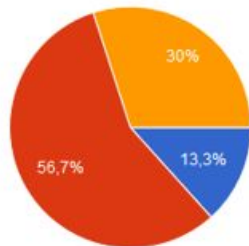**Regarding choices for the question, check if you encountered**

☐ It was easy to answer the question by an elimination of ridiculous answers

☐ I was deceived by other answers

☐ Question had more than one correct answer

**ODOSLAŤ**

**How difficult was the quiz?**

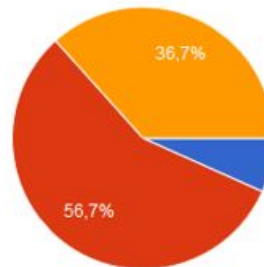RR — 30 answers

- Hard: 13,3%
- Medium: 56,7%
- Easy: 30%

**How difficult was the quiz?**

TR — 30 answers

- Hard
- Medium: 56,7%
- Easy: 36,7%

**How difficult was the quiz?**

RT — 23 answers

- Hard: 13%
- Medium: 60,9%
- Easy: 26,1%

**How difficult was the quiz?**

TT — 26 answers

- Hard: 23,1%
- Medium: 65,4%
- Easy: 11,5%
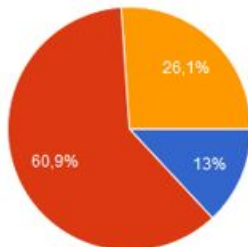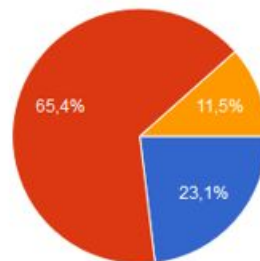
**TR = trained questions, random answers**

Question #2

| | Easy | Deceived | More than one correct |
|------|------|----------|----------------------|
| RR | 20 | 11 | 3 |
| TR | 27 | 9 | 1 |
| RT | 11 | 12 | 2 |
| TT | 17 | 12 | 4 |

**TR=trained** questions, **random** answers

22
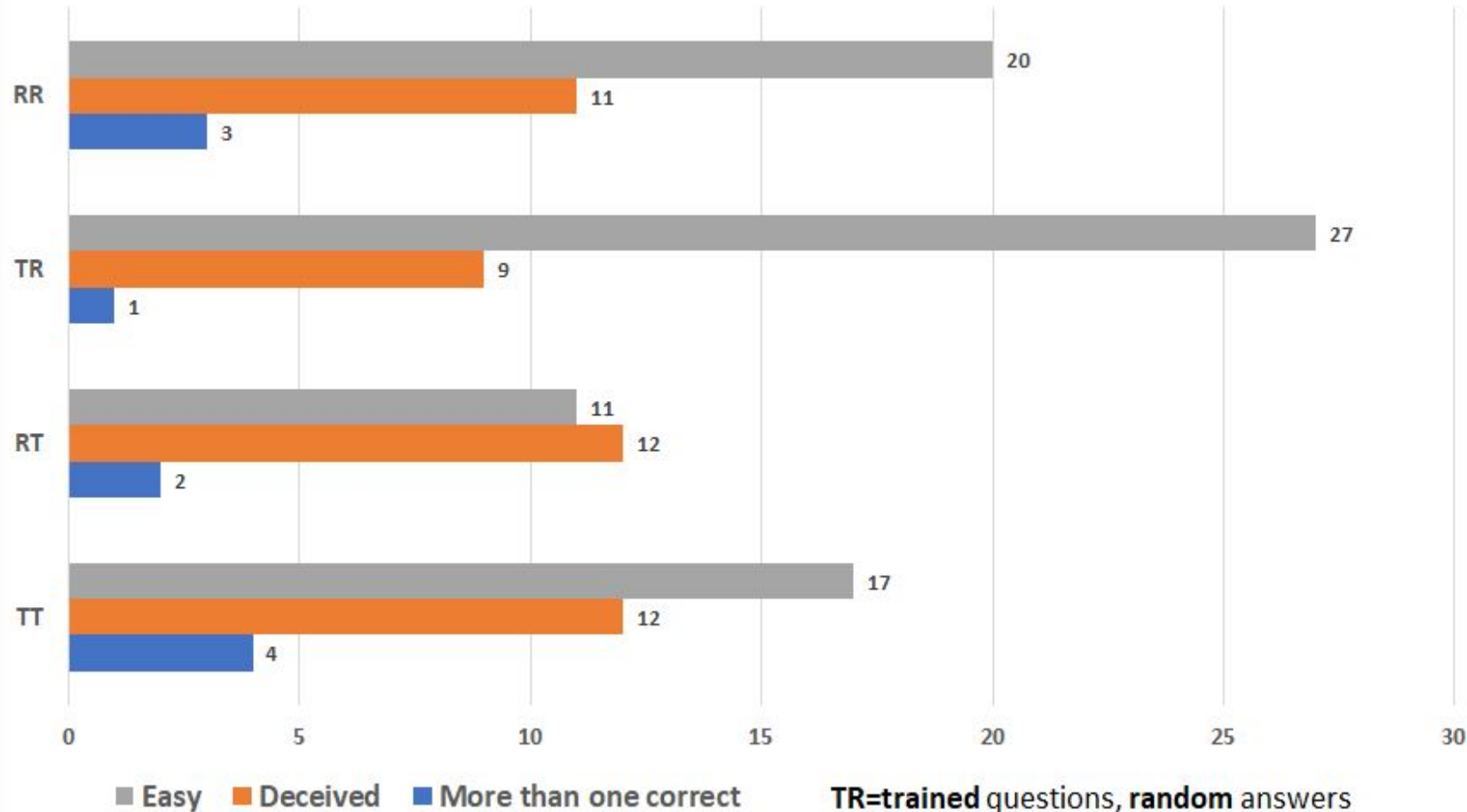
# Encountered challenges

- A lot of time spent on the plain application, not the recommender part
  - It was hard to agree on technologies (especially BE language) we are going to use, we tried "the middle ground" way
  - Wasn't too successful
- Prevent cheating
  - **Solution**: Sending correct answer to FE **after** the question was answered
  - Shuffle answers also on BE, so correct answer isn't always the first
- Finding big enough source of idioms
  - Selected source contained definitions which included idioms, we needed to scrape them manually
- Finding big enough corpus containing significant part of our idioms database
  - Problem to find an appropriate domain of dataset (not too technical, not too formal)

# Future work

- Improve recommendation of both questions and answers
    - Incorporate collaborative part to recommending questions
    - Adjust difficulties of idioms based on user answers
    - Use more features for answers generation + parameter tuning
- Include more types of questions
- Incorporate the system to existing English learning portal
- Improve our English by using the current system : )
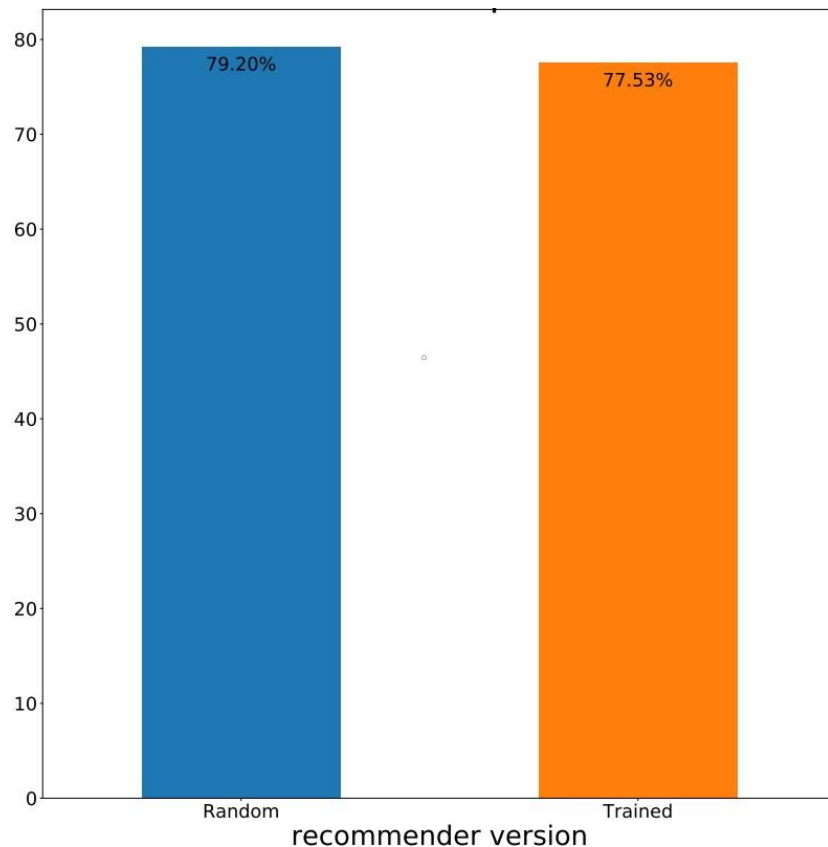
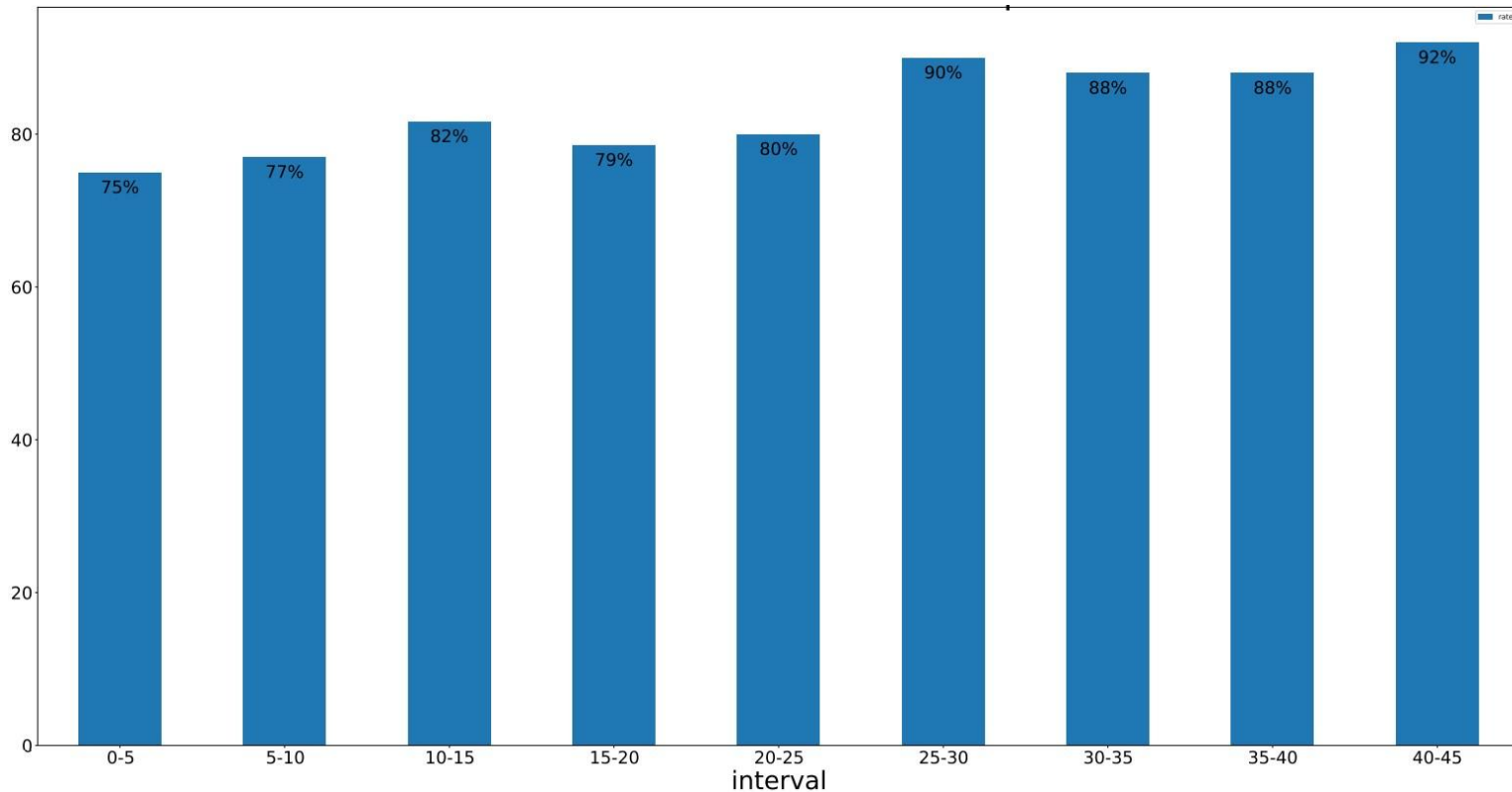# Thank you for your attention

# Other results

# Success rate - answers versions
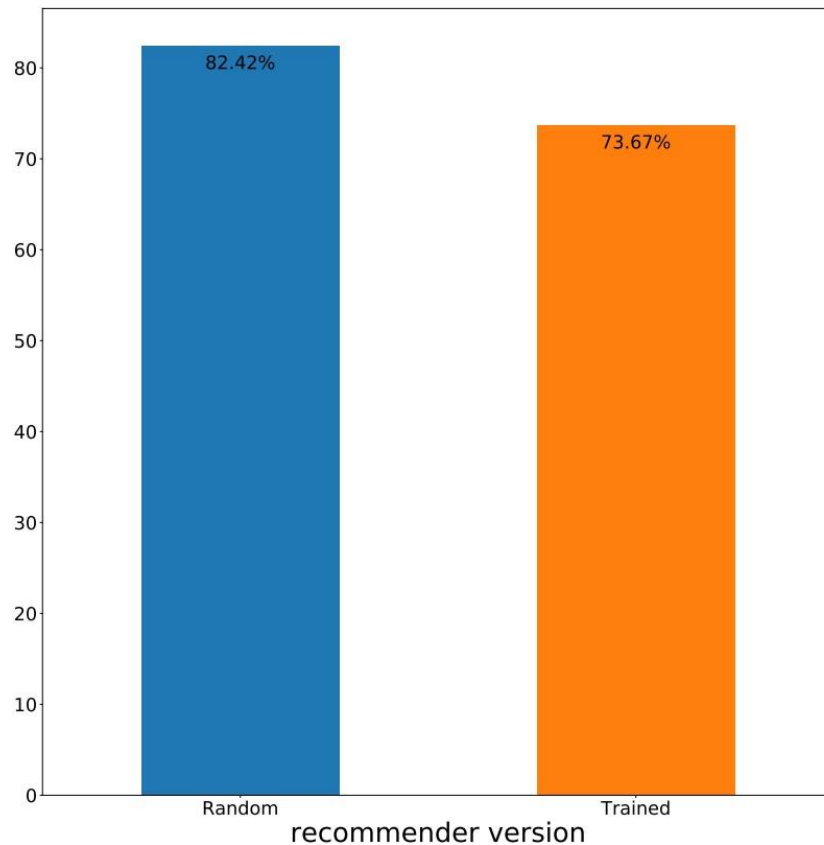
# Success rate - question versions

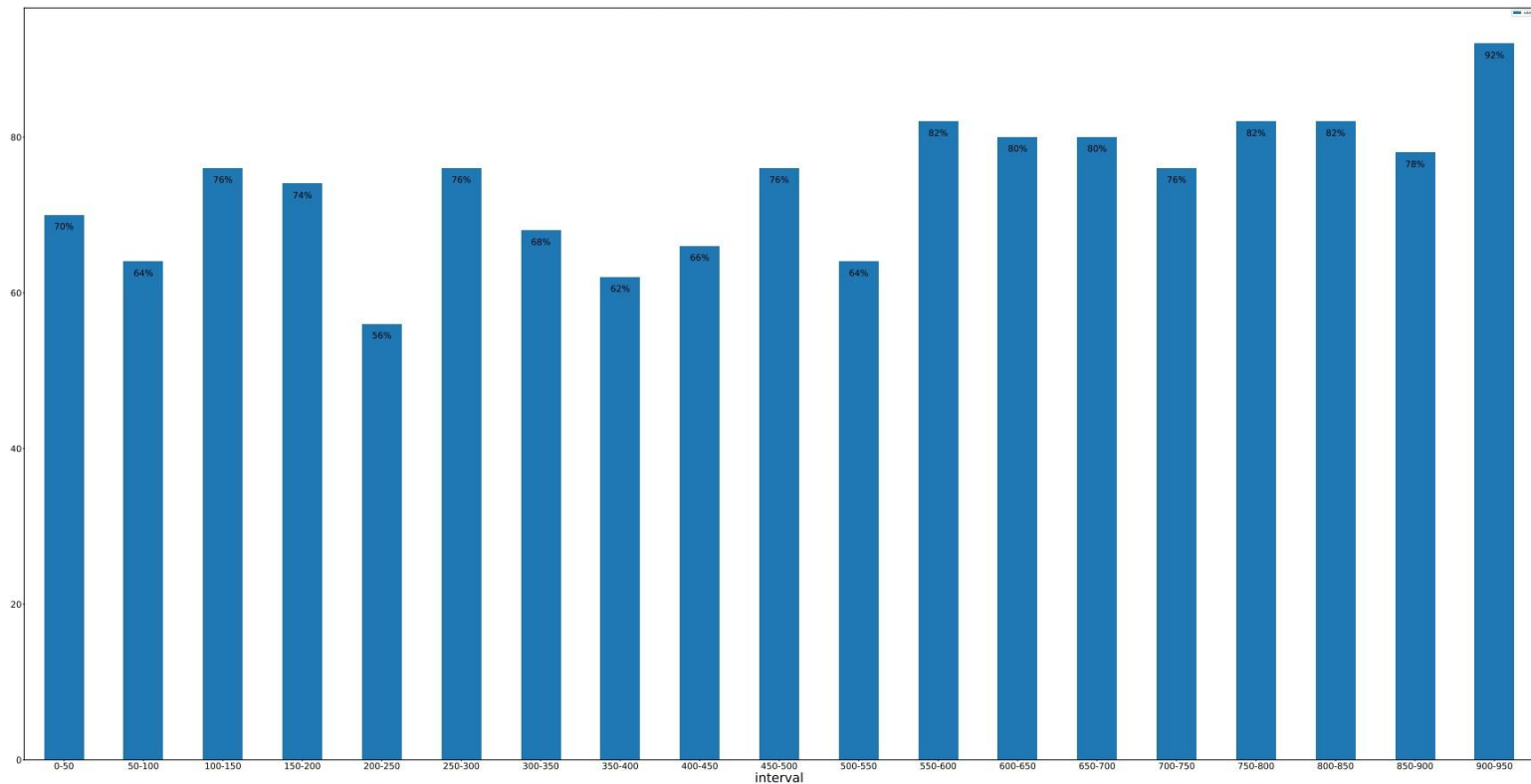# Success rate - R questions in intervals

# Success rate - T answers in intervals

# Success rate - R answers in intervals