

# Can we make learning foreign languages more personalized?

Mircea Lungu

Dec 2022, University of Lausanne



Roughly **1.2 billion** people worldwide are currently learning a foreign language.

<https://techcrunch.com/2017/03/05/3-habits-of-successful-language-learners/>

**And your's truly is one of them :)**

In fact I have been one for the last 20 years



**Blauzungekrankheit!**



**After a while textbooks and apps like Duolingo become boring**

**Most language materials are designed for the *average learner***

**No Freedom =>  
Demotivation**

**There is no  
such thing as  
the *average*  
user**

## When U.S. air force discovered the flaw of averages

TR

By Todd Rose

▲ Sat., Jan. 16, 2016 | 11 min. read

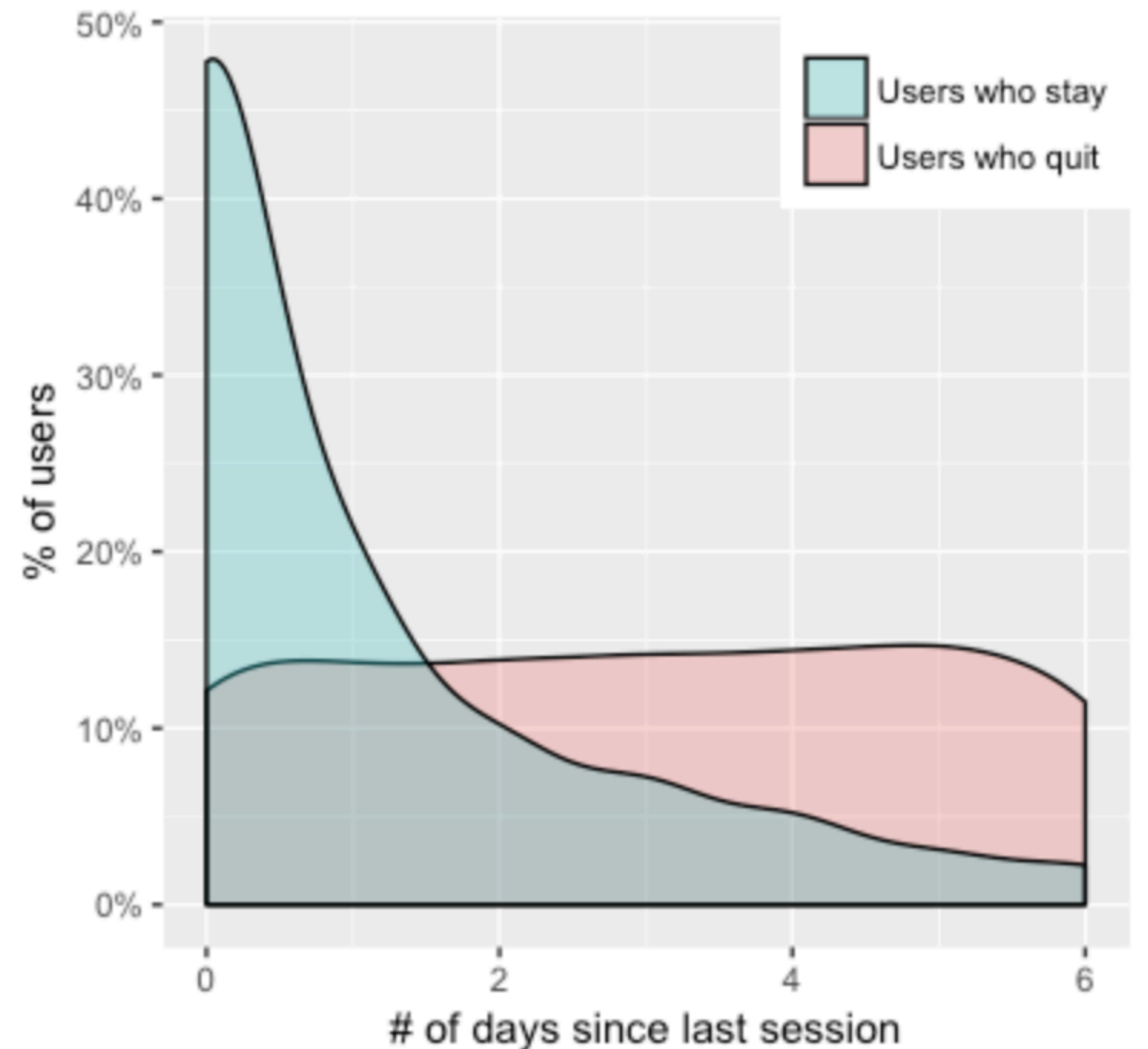


<https://www.thestar.com/news/insight/2016/01/16/when-us-air-force-discovered-the-flaw-of-averages.html>

# Best predictor of language learning? Motivation.

Suzanne Hidi and Judith M Harackiewicz.  
**Motivating the academically unmotivated:** A critical issue for the 21st century. *Review of educational research* 70, 2 (2000), 151–179.

William G Brozo, Gerry Shiel, and Keith Topping. **Engagement in reading: Lessons learned from three PISA countries.** *Journal of Adolescent & Adult Literacy*



# One of the best tools for learning? **Extensive Reading.**

1. **Students read a lot and read often.**
2. **There is a wide variety of text types and topics to choose from.**
3. **The texts are not just interesting: they are engaging/compelling.**
4. **Students choose what to read.**
5. **Reading purposes focus on: pleasure, information and general understanding.**
6. Reading is its own reward.
7. There are no tests, no exercises, no questions and no dictionaries.
8. **Materials are within the language competence of the students.**
9. Reading is individual, and silent.
10. Speed is faster, not deliberate and slow.

Willy A. Renandya. **The Power of Extensive Reading.**

Timothy Bell. **Extensive Reading: Why? and How?** *The Internet TESL Journal* 4

C McCarthy. **Reading theory as a microcosm of the four skills.** *The Internet TESL Journal* 5

Browser tabs: NZZ – Neue Zürcher Zeitung | x

Address bar: https://www.nzz.ch

Logo: **Neue Zürcher Zeitung**



**KOMMENTAR**

## Intrigen, Selbstherrlichkeit und Absurditäten: Das Parlament wird von der Schweizer Justiz hinters Licht geführt

Ein umstrittener Bundesgerichtspräsident, skurrile Geschichten am Bundesstrafgericht und Fehden am Bundesverwaltungsgericht – die Schweizer Justiz auf Abwegen.

Katharina Fontana 12.12.2022

Browser tabs: Issues - it-university | x | NZZ – Neue Zürcher Zeitung | x | Der Bund | Ihre Tradition | x

Address bar: https://www.derbund.ch

Logo: **Der Bund**

Navigation: Bern Aus den Regionen Schweiz Ukraine-Krieg



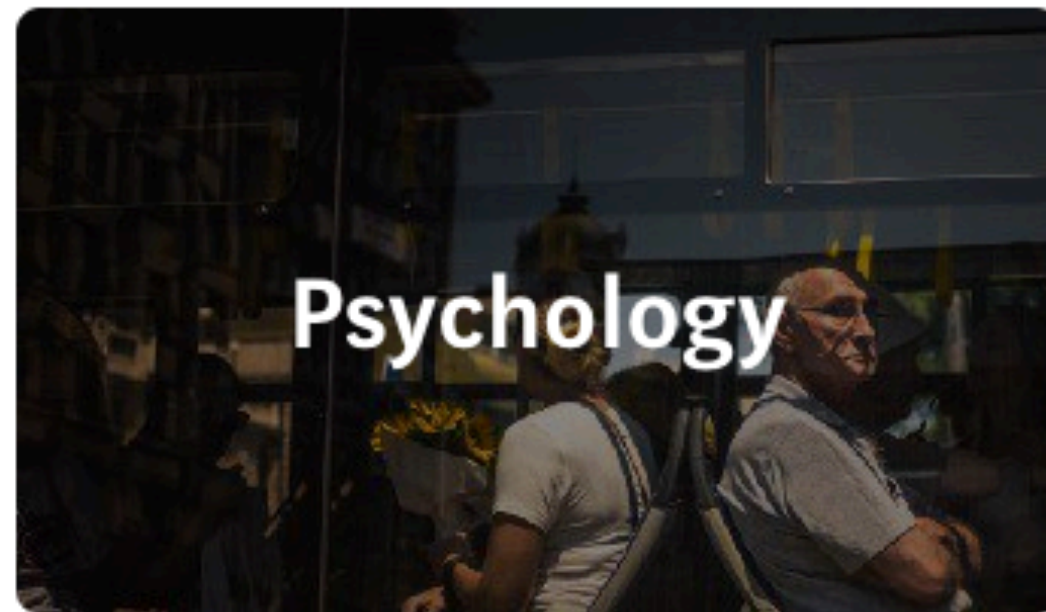
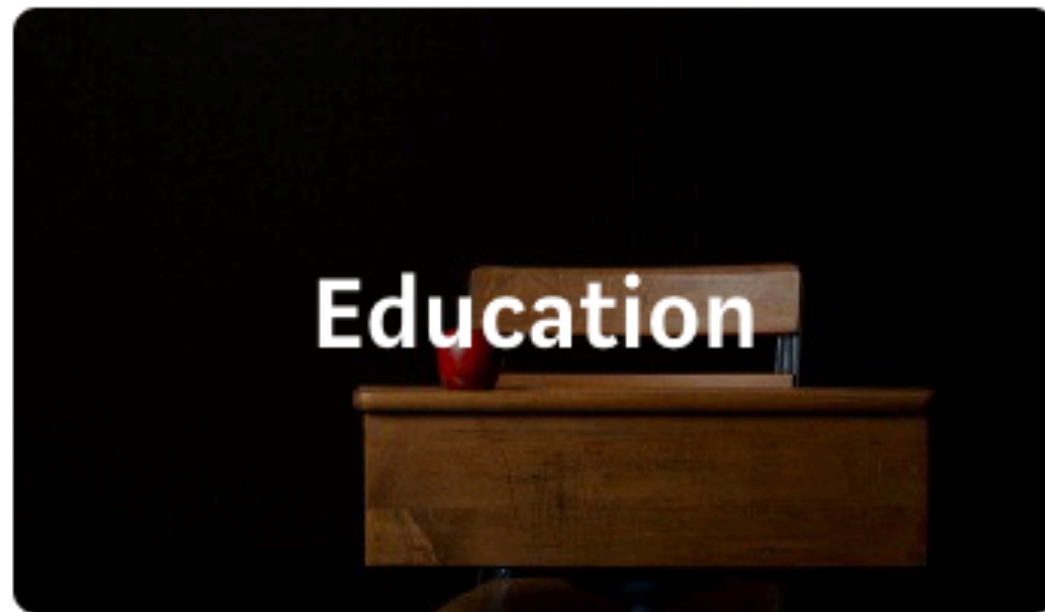
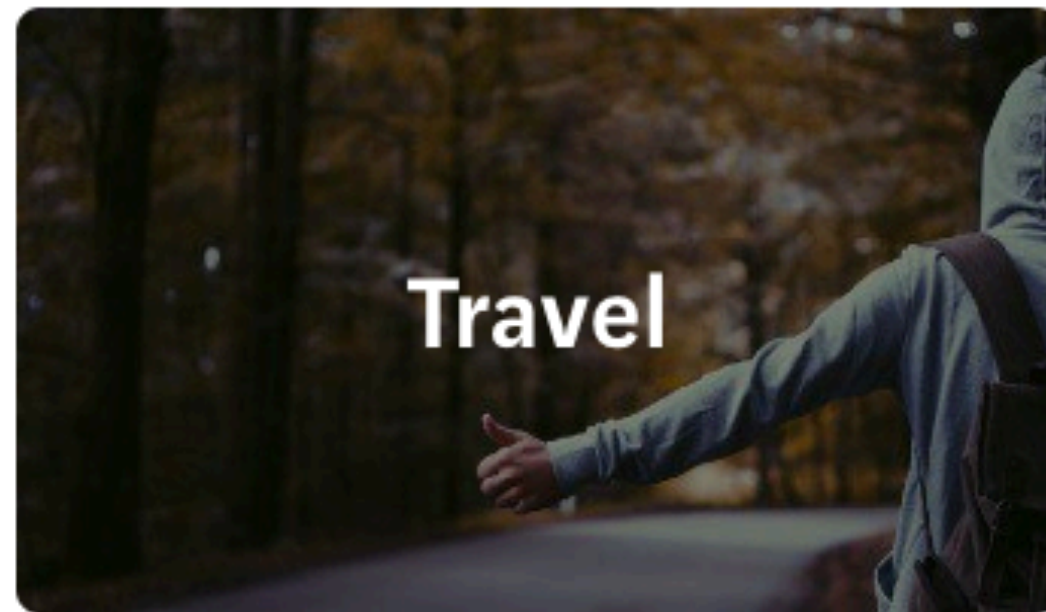
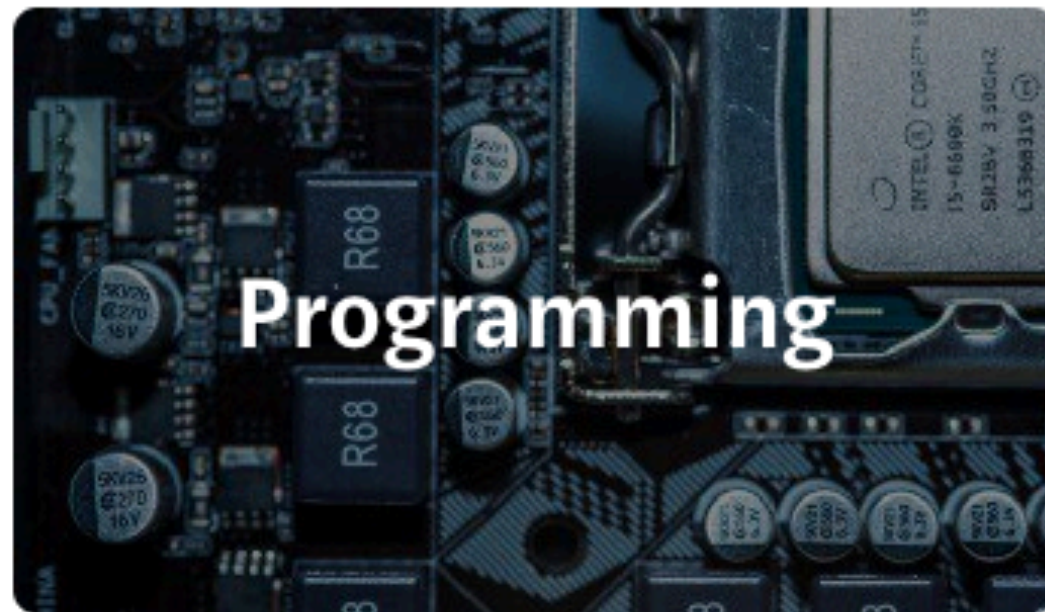
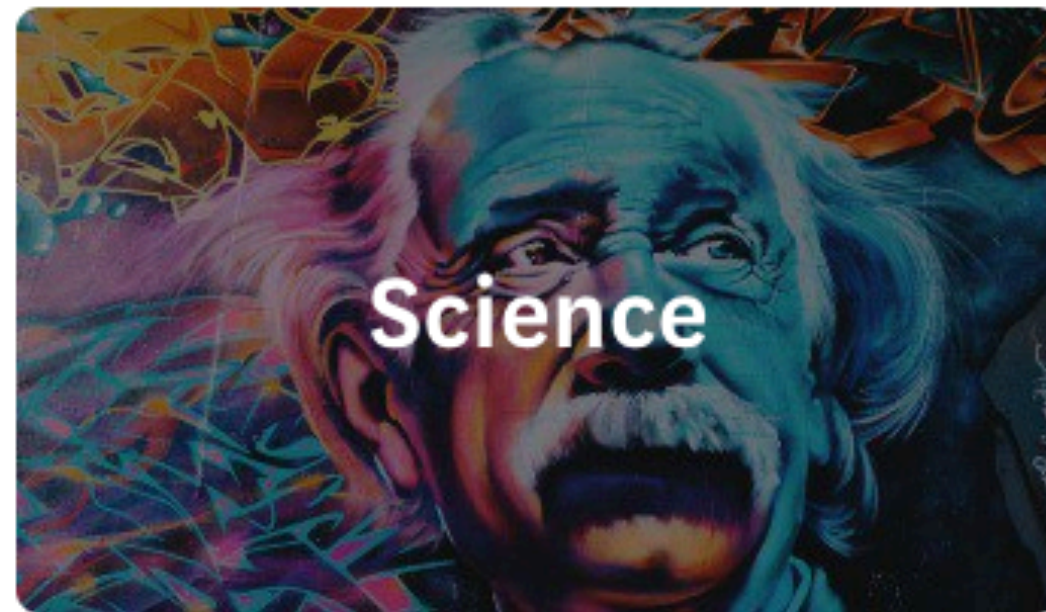
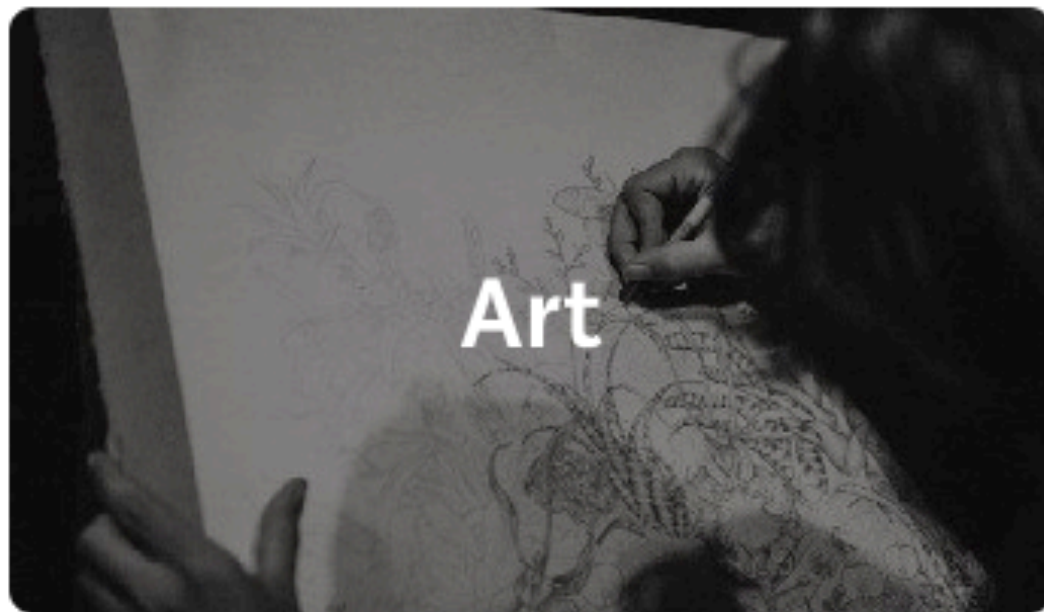
**Abo** **Besuch in Gerbers Fondue-Fabrik**

## Heimat, Natur und Eigenständigkeit – vereint in 400 Gramm Käse

Warum ist das Schweizer Nationalgericht so beliebt? Und wie gut ist die Mischung vom Fließband wirklich? Besuch

**No Competence =>  
Demoralization**

**Interesting texts  
exist on the net  
for everybody's  
interests and level**





RQ: Can we use the web as a  
**Personalized Textbook** for  
intermediate language learners?

# **Apropos of personalization...**

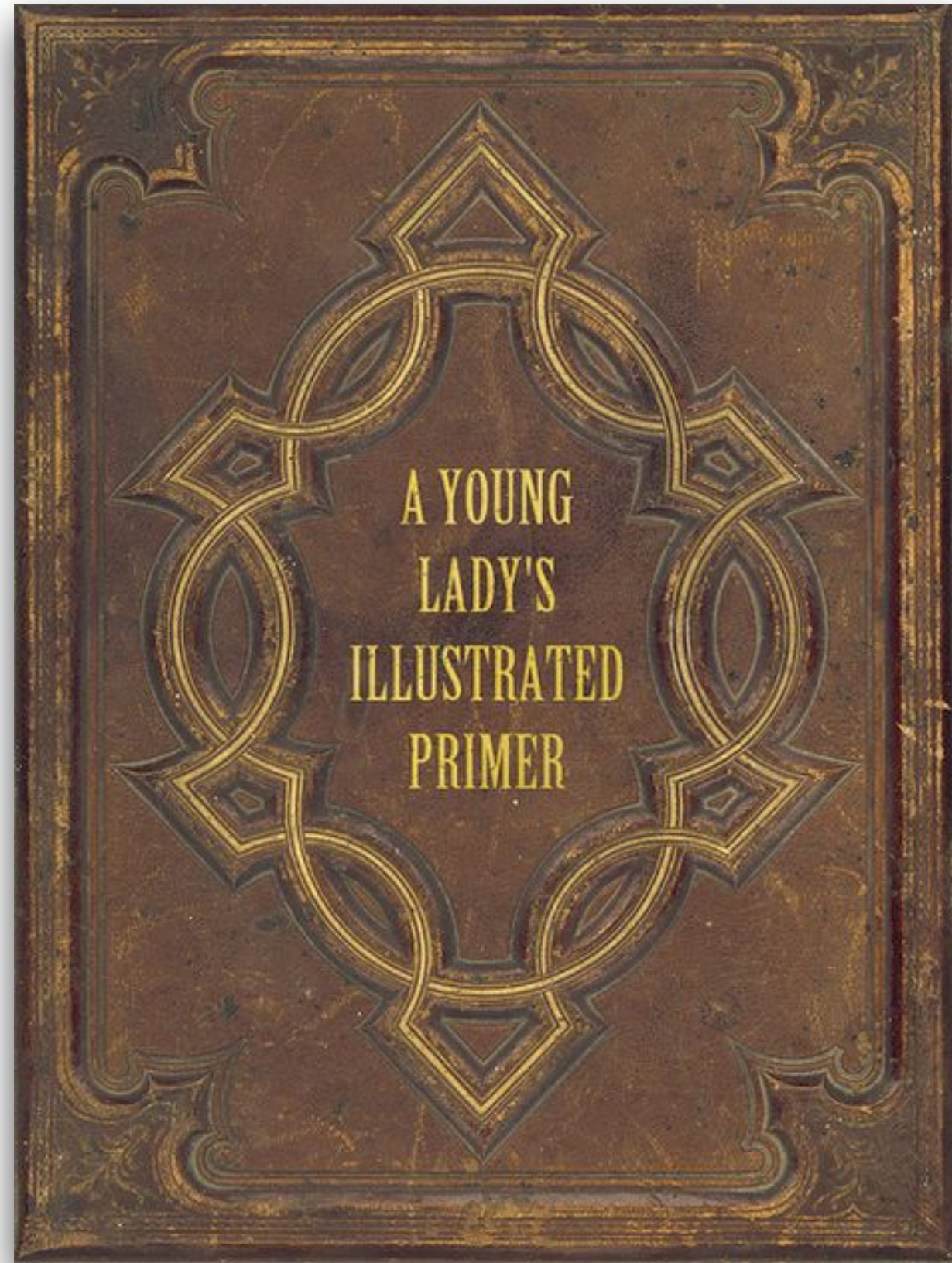
Who else uses it and why?

**WEAPONS OF**



**MASS DISTRACTION**

**Use **personalization** to motivate us to stay online**



**... fully interactive, able to teach reading, science, history and martial arts (among other things) while adapting to the current need of the reader.**

**Neal Stephenson, The Diamond Age  
or A Young Lady's Illustrated Primer**



**zeequu.org**

**An MVP for a**  
***PERSONALIZED***  
**Language**  
**Textbook**

Note: following are some of our solutions of solving these previously identified challenges; many are still work in progress.our solution is more of an MVP at the moment; each challenge is addressed as basic as possible so we can start testing and learning more.

# A **PERSONALIZED** Language Textbook Should

1. Enable Learners to **Read What They Like, Easily**



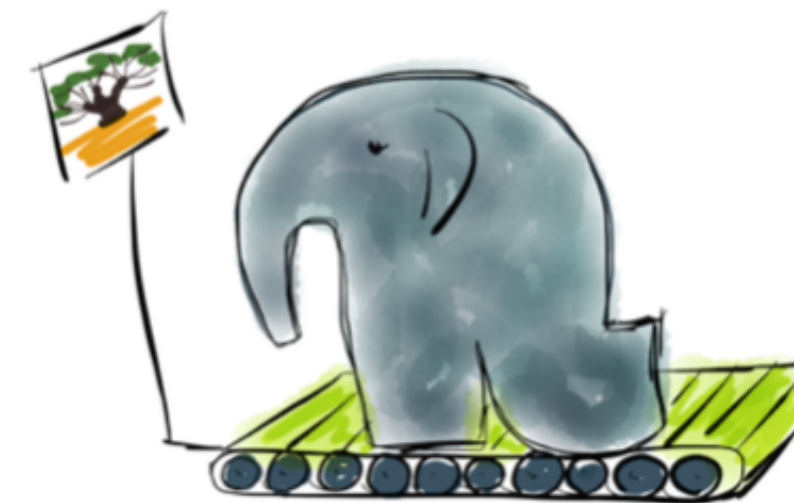
HCI & ML Challenge

2. **Generate Exercises Based on Past Readings**



Algorithmic Challenge

3. Provide **Insights Into Activity and Progress**



HCI & Algorithmic Challenge

# 1. Enable Learners to Read What They Like, Easily

- a. Support individual topic preferences
- b. Provide difficulty estimation
- c. Offer easy in-text translations



Read **What They Like**, Easily

# Support individual topic preferences

Interests

opt-in

Business (+) Save

Culture Food Health

Music Politics Science

Sport Technology

Travel World

Non-interests

opt-out

Business (+) Save

Culture Food Health

Music Politics Sport

Travel World Trump

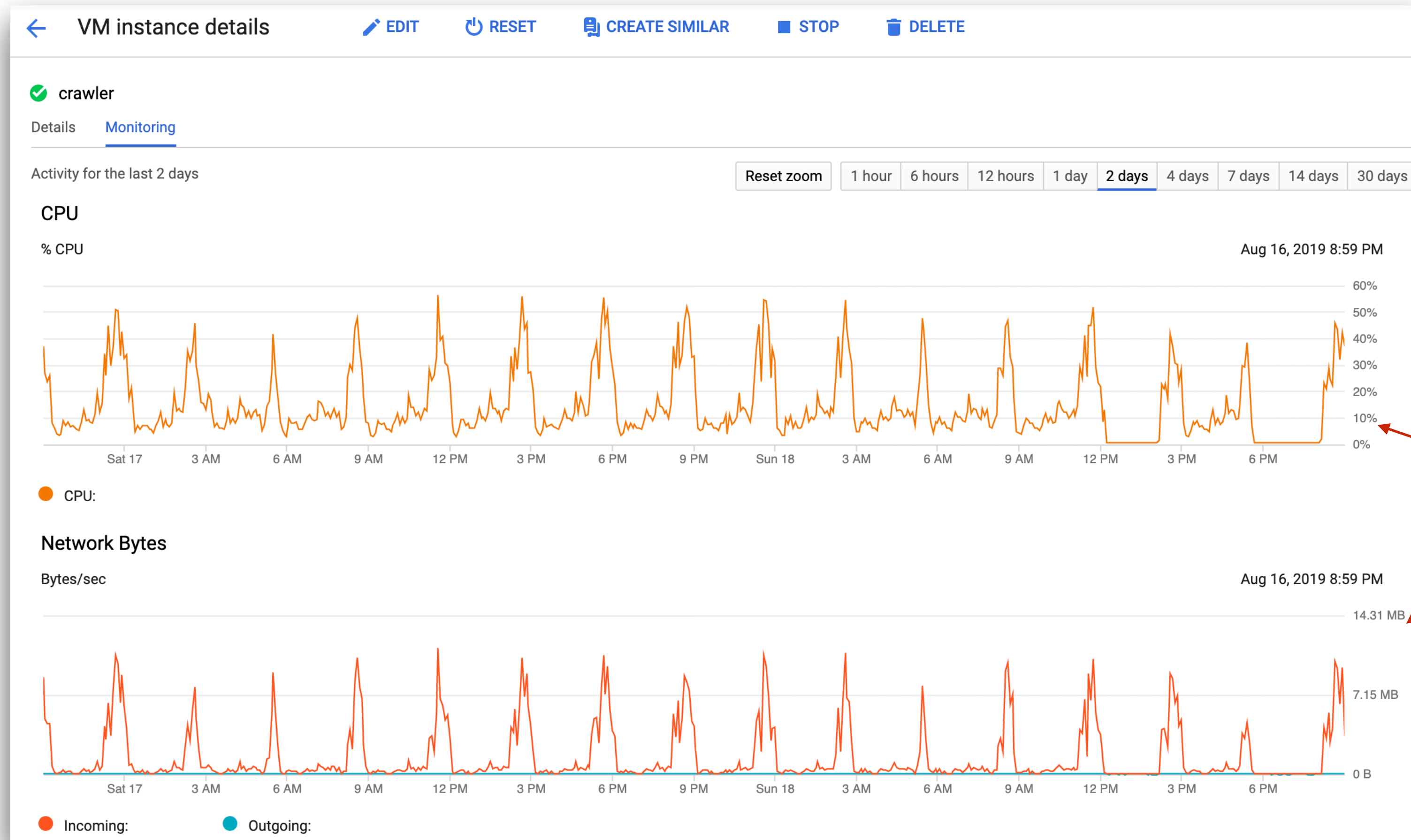
Musk

The screenshot shows the Zeeguu website interface. The browser address bar displays "https://www.zeeguu.org/articles". The page title is "Texts". There are navigation tabs: "Recommended", "Classroom", "Bookmarked", and "My Texts". Below the tabs, there are buttons for "Interests" and "Non-interests", and a search bar labeled "Search all articles". A sidebar menu on the left contains: "Texts", "Words", "Exercises", "Statistics", "Settings", "Feedback", and "Logout". The main content area displays three article recommendations, each with a title, a short description, a date, a topic tag, and a rating. The first article is "Bébé zéro déchet : c'est bon pour lui, la planète et votre porte-monnaie !" with a rating of 4.2. The second is "La fin de vie, le dernier combat des francs-maçons en France" with a rating of 4.4. The third is "C'est quoi, le métier d'astronaute ?" with a rating of 3.3.

Read **What They Like**, Easily

# Support individual topic preferences

**Under the Hood**



**Multiple Times a Day  
Find New Reading  
Materials**

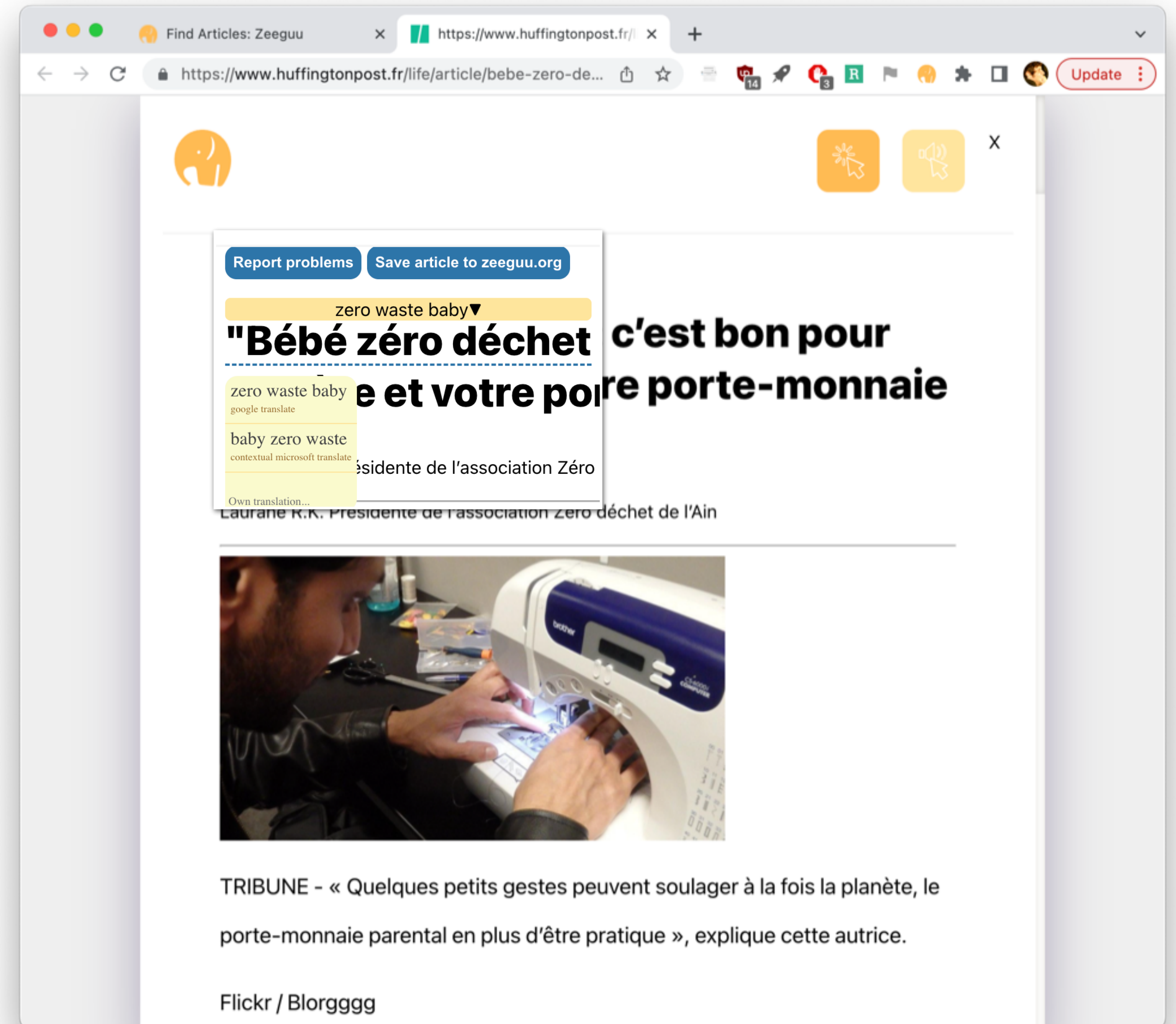
# Estimate difficulty & Propose Readings

The screenshot shows a web browser window with the URL <https://www.zeeguu.org/articles>. The page features a sidebar on the left with navigation options: Texts, Words, Exercises, Statistics, Settings, Feedback, and Logout. The main content area displays a list of articles, each with a title, author information, date, category, and a rating in a circle.

Title	Author	Date	Category	Views	Rating
Bébé zéro déchet : c'est bon pour lui, la planète et votre porte-monnaie !	Flickr / Blorgggg Flickr / Blorgggg ENVIRONNEMENT	(3 days ago)	Science	1706	4.2
La fin de vie, le dernier combat des francs-maçons en France	Emmanuel Macron	(3 days ago)	Science	226	4.4
C'est quoi, le métier d'astronaute ?	Cinq nouveaux astronautes	(4 days ago)	Science	311	3.3
Pourquoi ChatGPT a moins d'humour que Siri ?	dowell / Getty Images	(4 days ago)	Science	584	3.3

- Fleisch-Kinkaid remapped on a 1 – 10 scale
- Work in Progress with Michalis and Yash to add a supervised learning approach that would rate CEFR level

# Support reading with easy translation & pronunciation



- Approx. 1/7 translations => dropdown
- The teacher who liked the imperfection

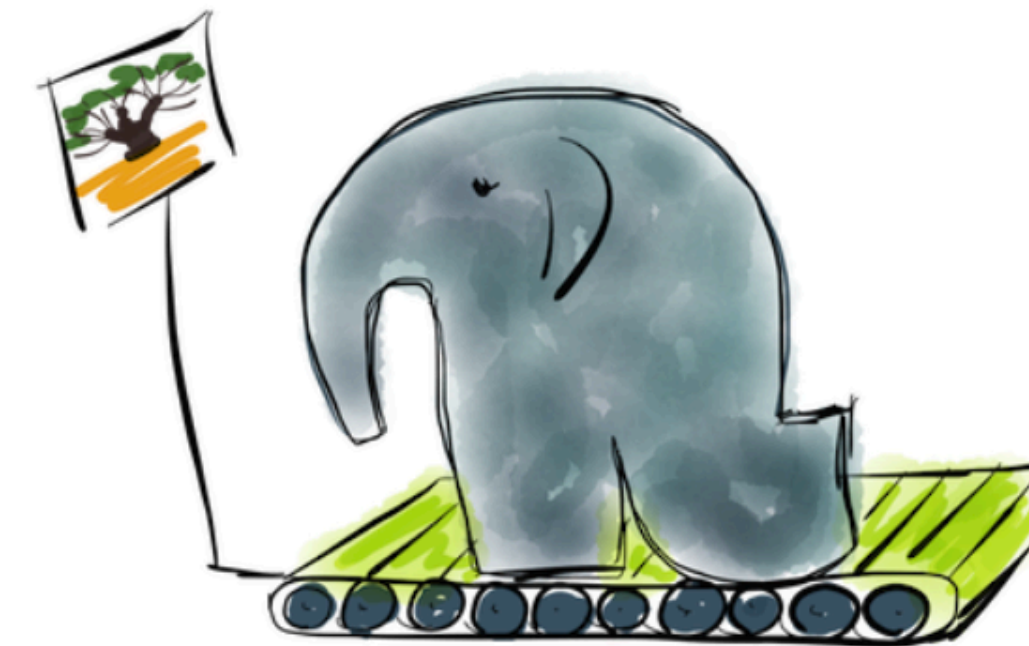
# Demo...



[zeequu.org](http://zeequu.org)

## 2. **Generate** Personalized Vocabulary Exercises Based on Past Readings

- a. **Decide which words to prioritize**
- b. **Consider multiple exercise types**



# Multiple exercise types

- Users want multiple exercise types so they are not bored
- Different exercise types have different difficulties
- Some contexts are too long: don't practice them... problem with teachers who don't understand!

The screenshot shows a 'Triple Choice' exercise. At the top, a yellow header contains the text 'Triple Choice'. Below the header, a paragraph of German text reads: 'Ich habe den Eindruck, das hat damit zu tun, dass wir eher einer Realpolitik als einer Symbolpolitik \_\_\_\_\_ .'. At the bottom of the interface, three buttons are visible: 'angehen', 'Fürstel', and 'verpflichtet sind'.

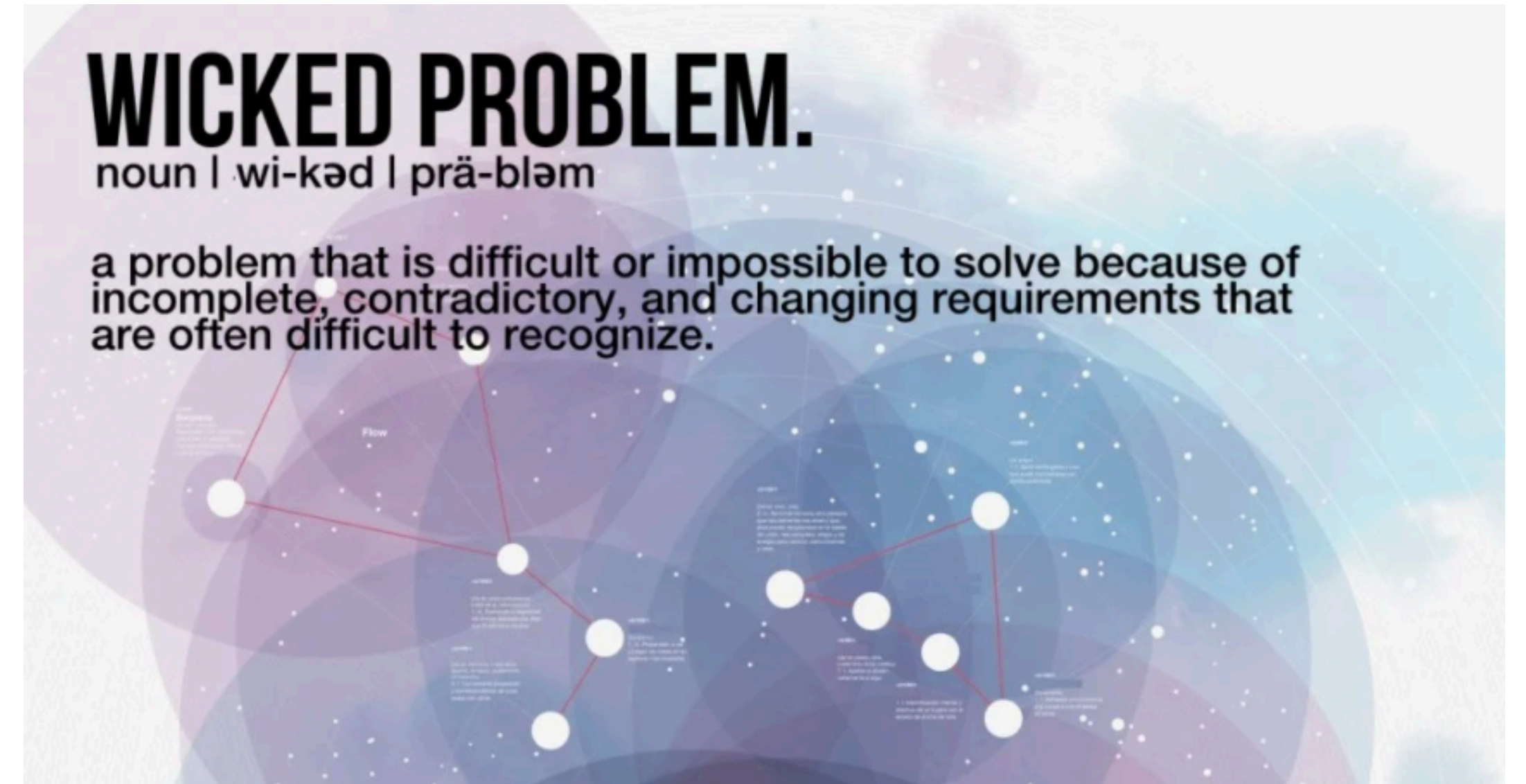
The screenshot shows a 'Match' exercise. At the top, a green header contains the text 'Match'. Below the header, there are two columns of words. The left column contains 'umgestellt würde' and 'verpflichtet sind'. The right column contains 'grouchy', 'Would be reversed', and 'Are required'. At the bottom of the interface, a grey bar contains a question mark icon.

The screenshot shows a 'Recognize & Type' exercise. At the top, an orange header contains the text 'Recognize & Type'. Below the header, a paragraph of German text reads: 'Die beiden anderen grossen Temporärunternehmen, Manpower und Kelly Services, planen dagegen gegenwärtig keine vergleichbaren Angebote.'. At the bottom of the interface, there is a grey bar with a question mark icon, a text input field with the placeholder 'Type or click a word', and a green 'Check' button.

# Decide which words to prioritize

## Factors to consider

- frequency
- already in learning
- recently looked up
- the allocated time
- spaced-repetition requirements
- user provided priority
- reaction times
- user availability



# 3. Provide **Insights Into** **Activity & Progress**

- a. **Feedback to Learner**
- b. **Insights for Teacher**



# Feedback for Learner



## Your Words




History | Starred | Ranked | **Learned**

Learned words are words that were correct in exercises on four different days or words that have been tagged to be 'Too easy'.

You have learned 22 words so far.

  **sladrer**  
gossip 

Correct on: 2022-11-21

  **Derfor**  
therefore 

Correct on: 2022-09-23

  **netop**  
exactly 

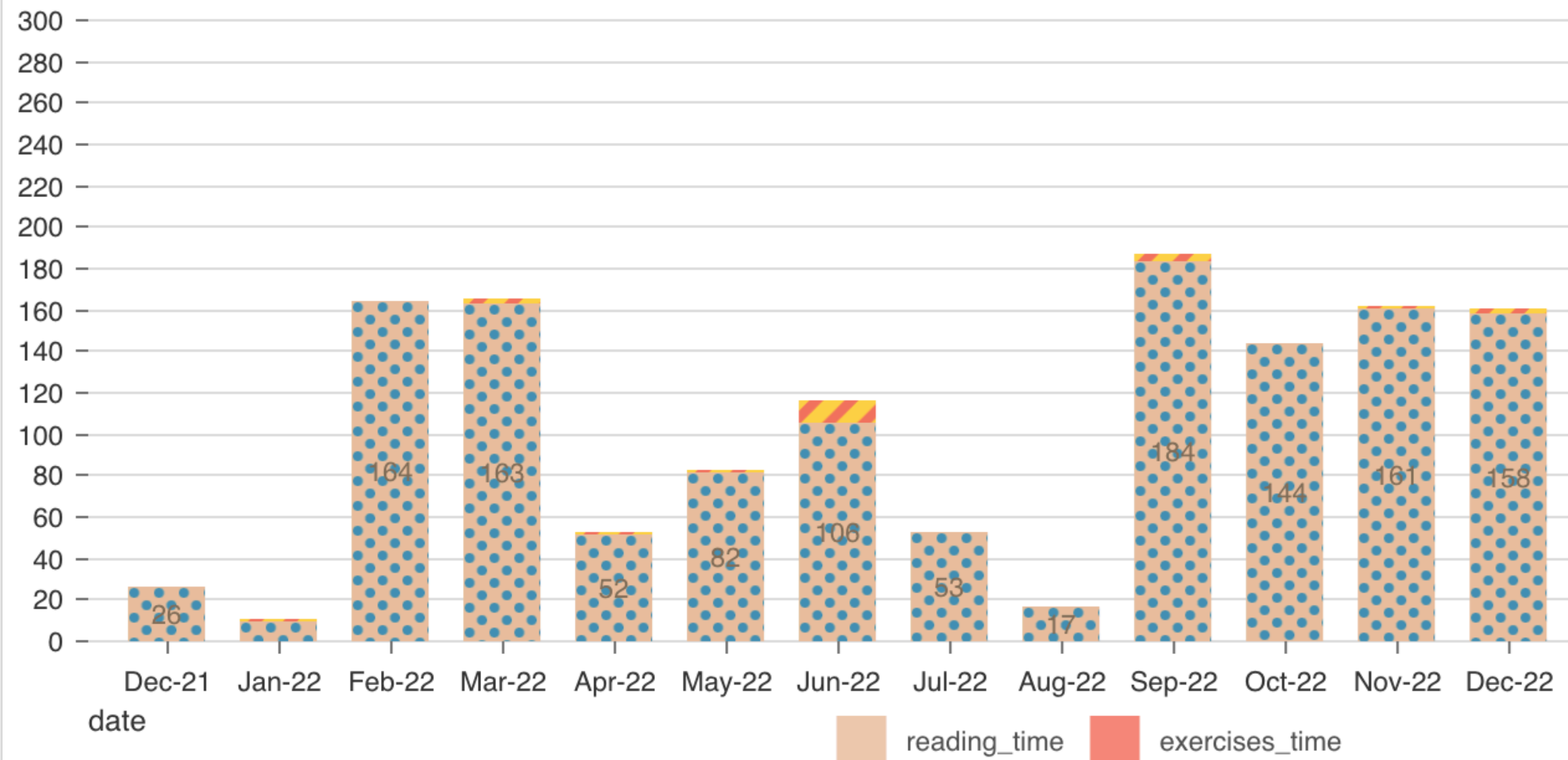
Correct on: 2022-06-27

  **ærligt**  
honestly 

Correct on: 2022-06-27

Activity Words

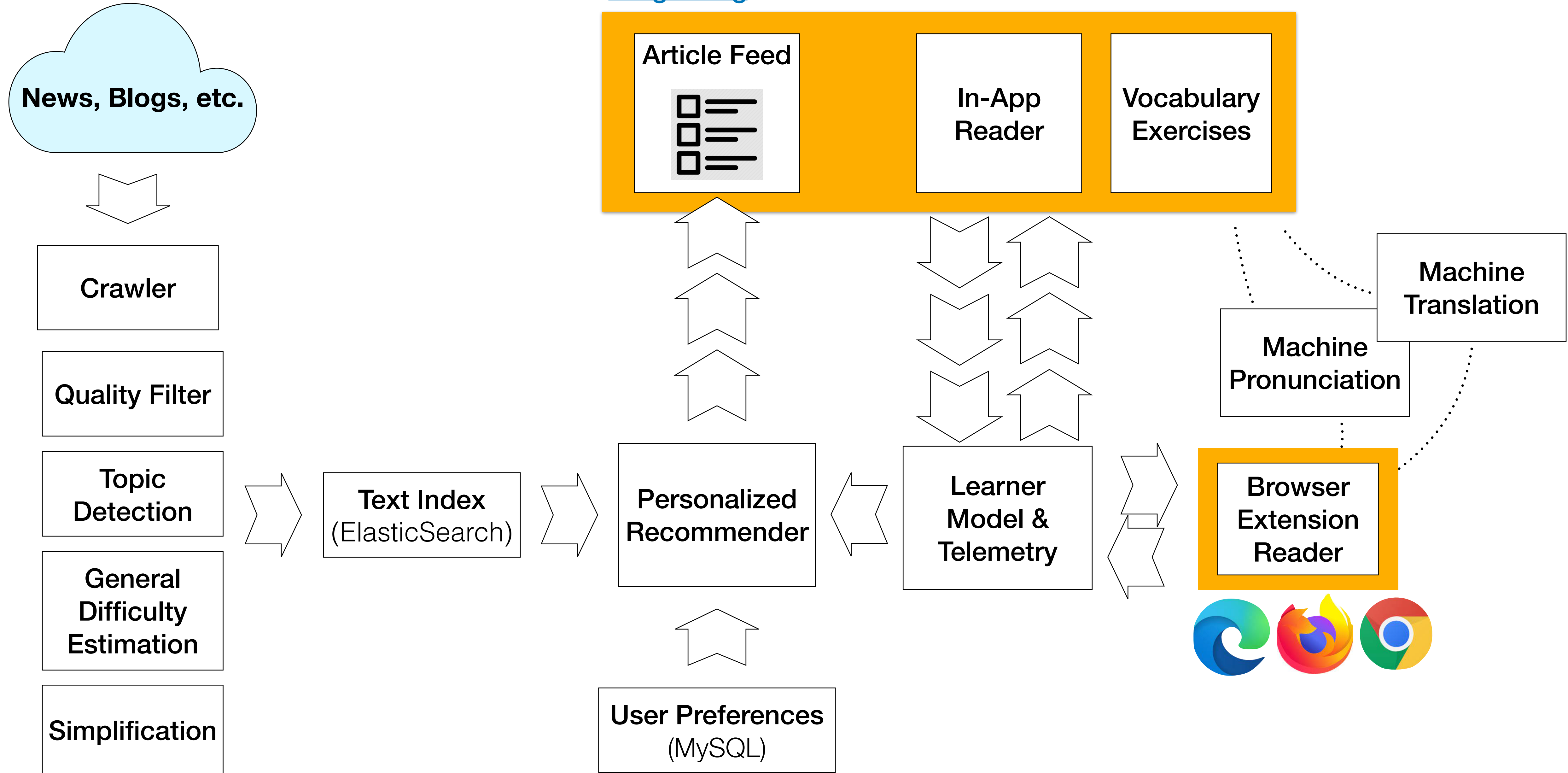
Your activity data for **Last 12 months** Time count shown in **minutes**



# (Bonus) Insights for Teacher

OD23A2+ Dutch				MANAGE ARTICLES	EDIT CLASS
Invite code: OD23A2+					
NAME	TIME SPENT ↓	CLASS NAME	ACTIVITY	Reading	Exercises
Aron	147h 19m	OD23A2+			
Maria	3h 25m	OD23A2+			
Erin	1h 24m	OD23A2+			
Vasile	1h 5m	OD23A2+			
Justin	0h 30m	OD23A2+			
Thijs	0h 15m	OD23A2+			
Anna-Maria	0h 10m	OD23A2+			

[zeeguu.org](http://zeeguu.org)



# **Some Learned Lessons**

**Learners  
benefit from  
personalization  
of content**

# 3 Weeks & 60 French Students

(bilingual Dutch-English)



**The Teacher's  
Perspective**

Presentation by Wim Gombert

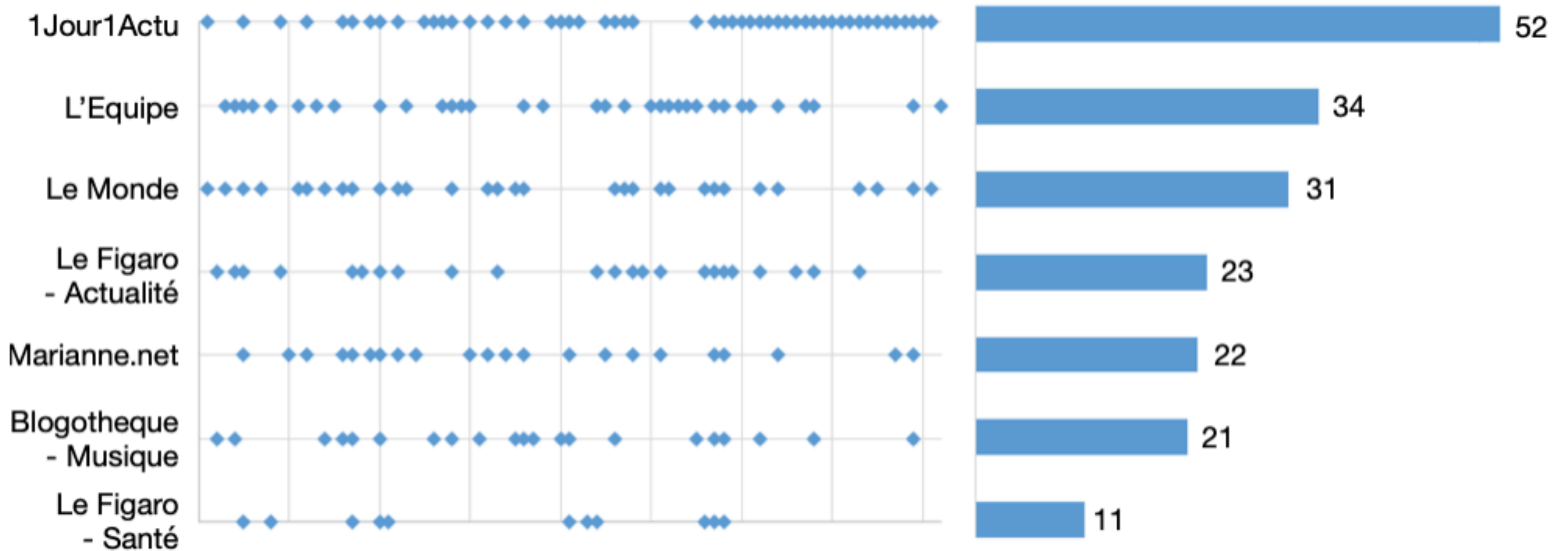
**Telemetry**

How do learners use the system?

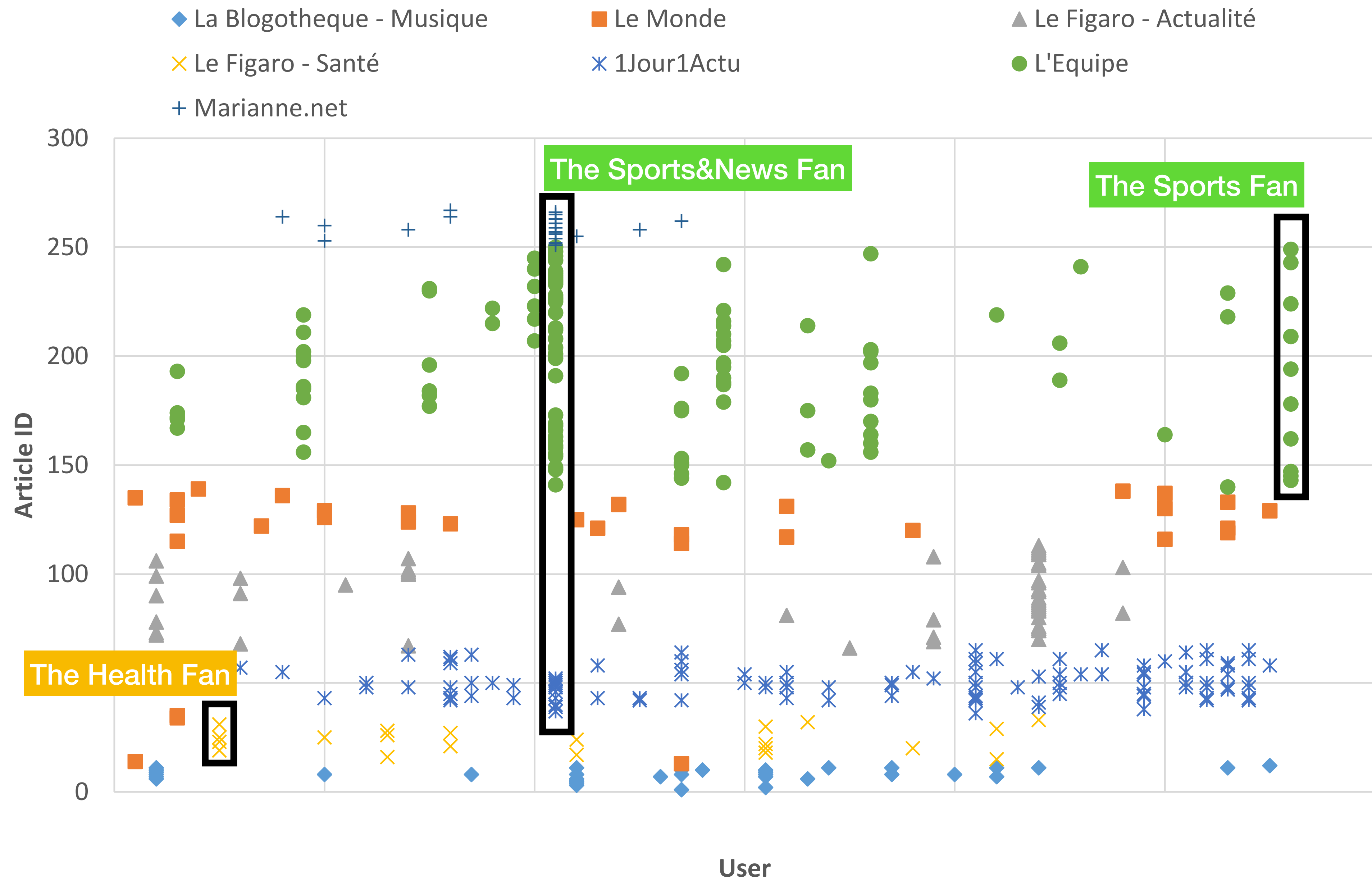
**Student  
Feedback**

My vocabulary truly is improving, but you do have to use it more than a few times. A very nice website, easy to use and with nice topics

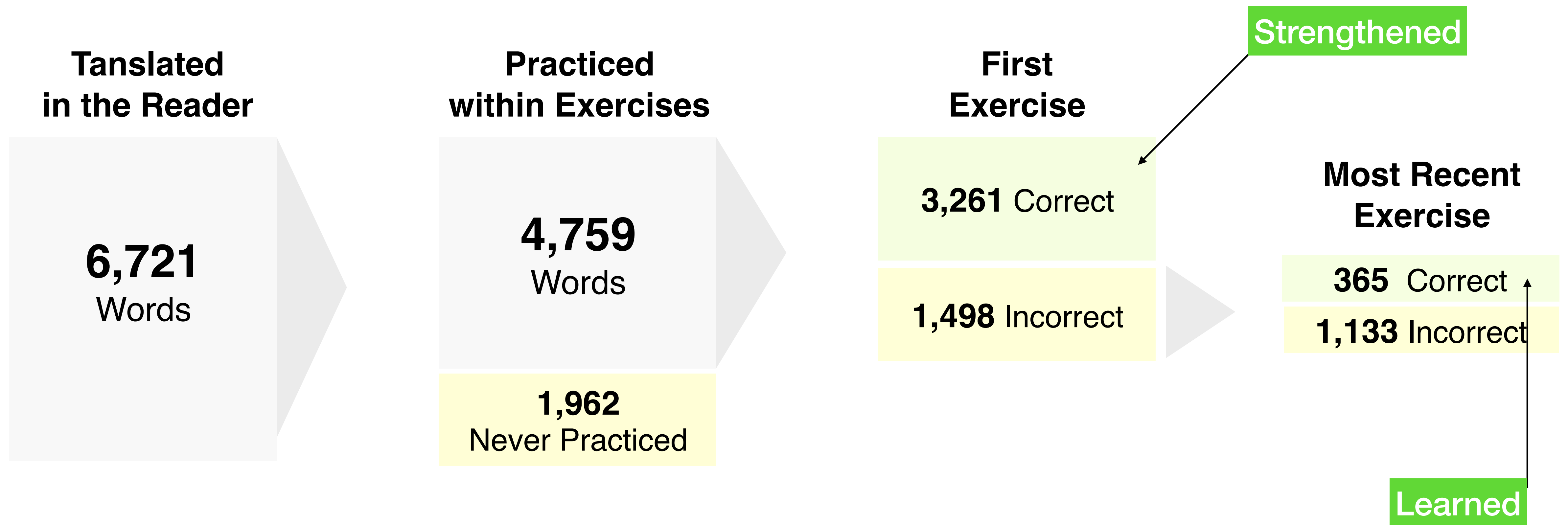
# Source Subscriptions



# Reader Types

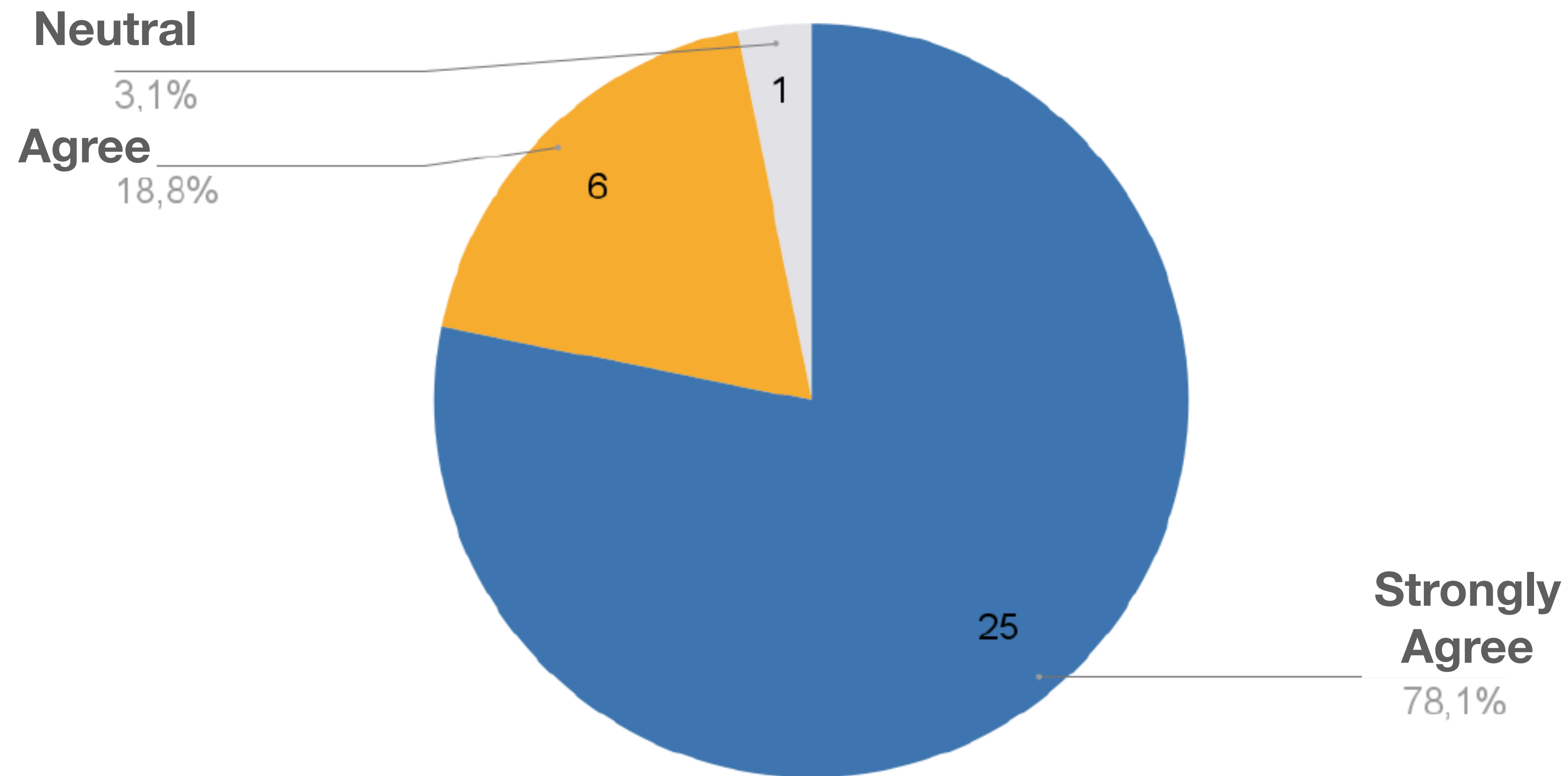


# Word Interactions In the Exercises



# A Recent (2022) Study with 32 participants

Do you like that you get to choose articles from the web yourself?



# Progress is not easy to quantify

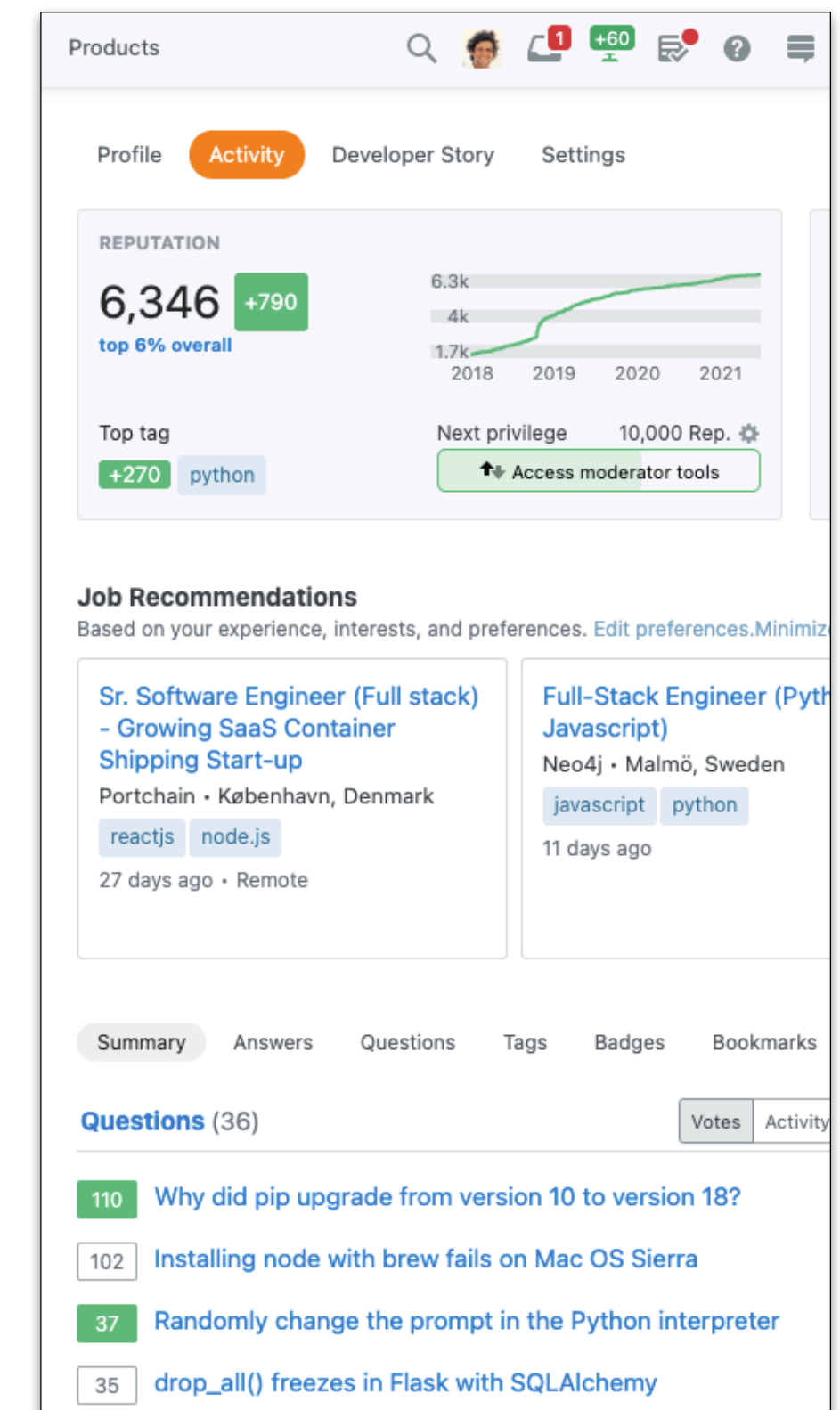
## Knowledge

- Not easy to quantify (what is a *learned word*?)
- Teachers seem happy with simple definitions

## Reading Time

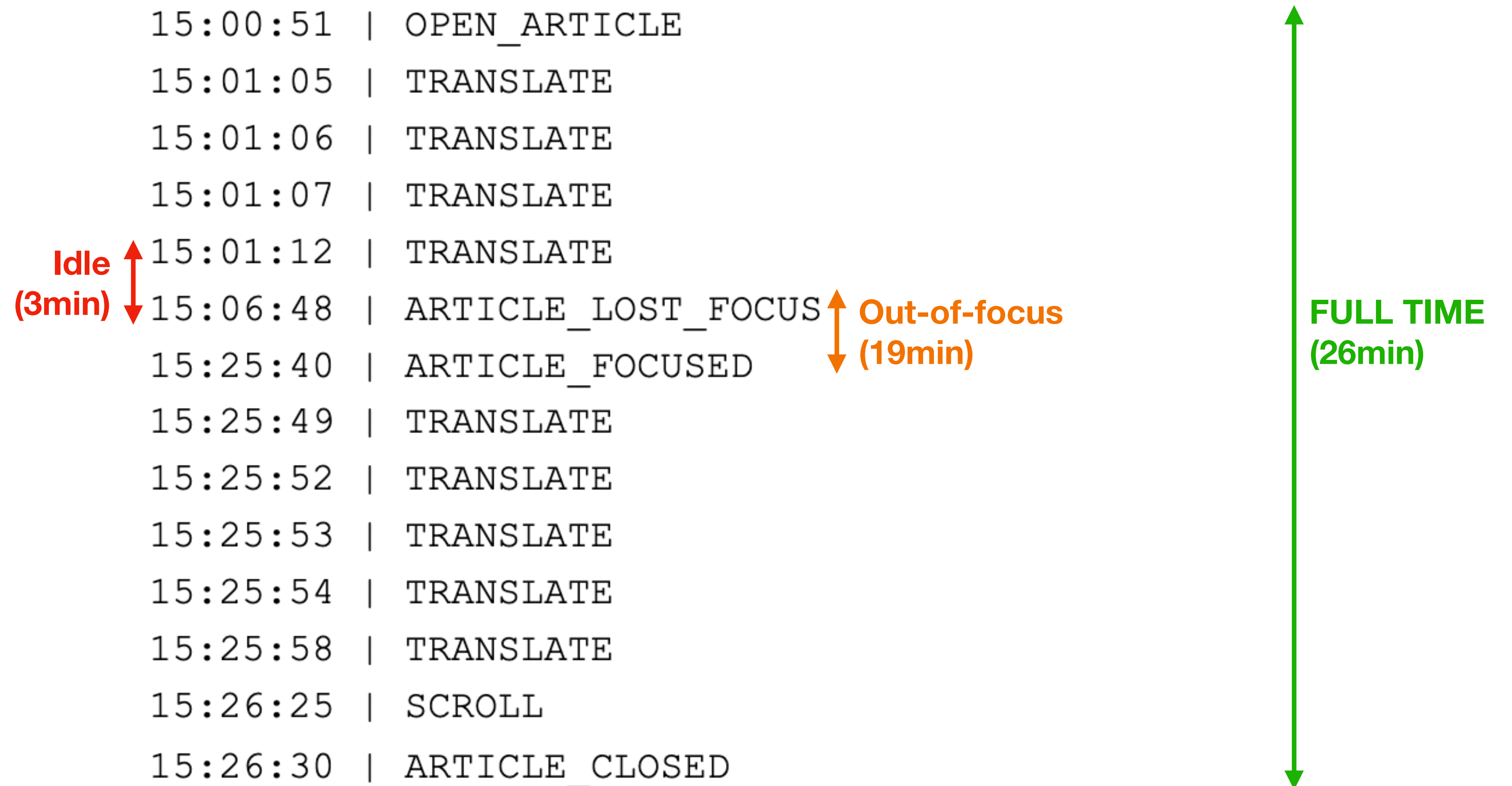
- Meaningful metric (as opposed e.g., “golden stars”, etc.)
- Could be also combined with gamification
- *Still surprises learners...*

A meaningful metric:  
reputation on StackOverflow



# Reading Time: **Common** User Behavior

**26m FULL-TIME / 4min FOCUSED-TIME!**



# Reading Time: **Extreme** User Behavior

Constantly losing focus over 2h

Date / Time	Article ID	Action	
2022-04-25 14:52:18	1910677	OPEN ARTICLE	
2022-04-25 14:52:19	1910677	ARTICLE LOST FOCUS	<b>Out-of-focus (8s)</b>
2022-04-25 14:53:27	1910677	ARTICLE FOCUSED	
2022-04-25 14:53:35	1910677	TRANSLATE TEXT	
2022-04-25 14:53:37	1910677	ARTICLE LOST FOCUS	<b>Out-of-focus (59s)</b>
2022-04-25 14:54:36	1910677	ARTICLE FOCUSED	
2022-04-25 14:54:40	1910677	TRANSLATE TEXT	
2022-04-25 14:54:44	1910677	ARTICLE LOST FOCUS	<b>Out-of-focus (19s)</b>
2022-04-25 14:55:03	1910677	ARTICLE FOCUSED	
2022-04-25 14:55:06	1910677	TRANSLATE TEXT	
2022-04-25 14:55:10	1910677	ARTICLE LOST FOCUS	<b>Out-of-focus (40s)</b>
2022-04-25 14:55:52	1910677	ARTICLE FOCUSED	
2022-04-25 14:55:58	1910677	TRANSLATE TEXT	

**90 events more... same pattern**

...

2022-04-25 16:48:48	1910677	ARTICLE FOCUSED	
2022-04-25 16:49:21	1910677	ARTICLE CLOSED	

# Audio Exercises Have a Positive Impact

- Controlled experiment (n=22)
- Increase self-reported
  - **perception of fun**
  - **likelihood of study**

In general learners are asking for more diversity  
In exercises.

How fun was your experience?

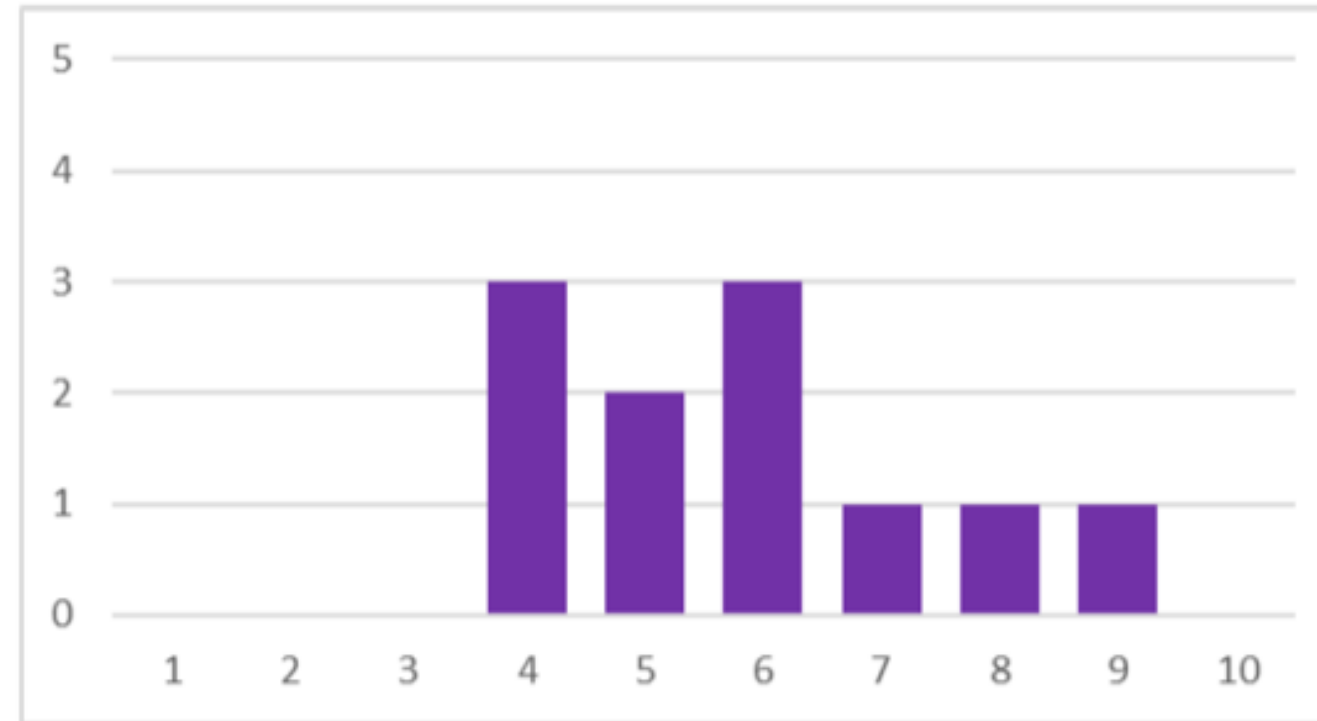


Figure 11. Frequency histogram of control group

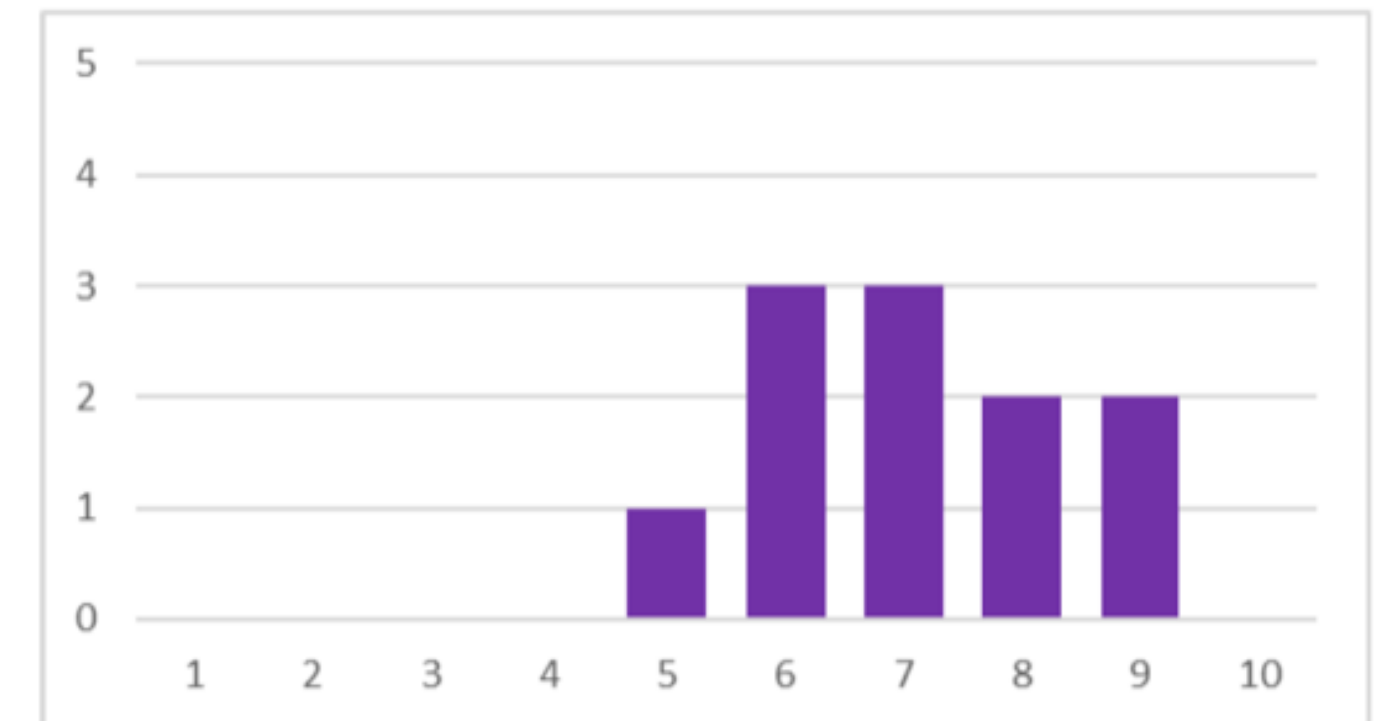
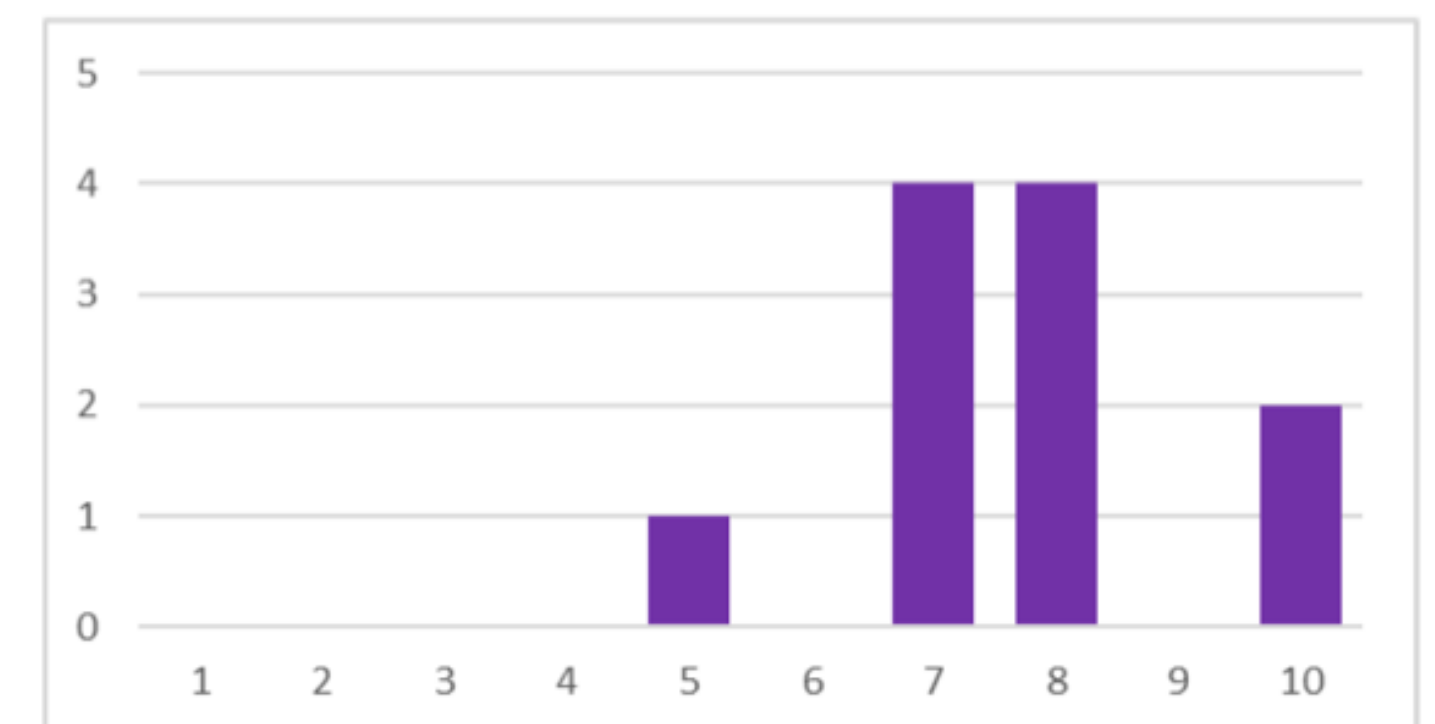
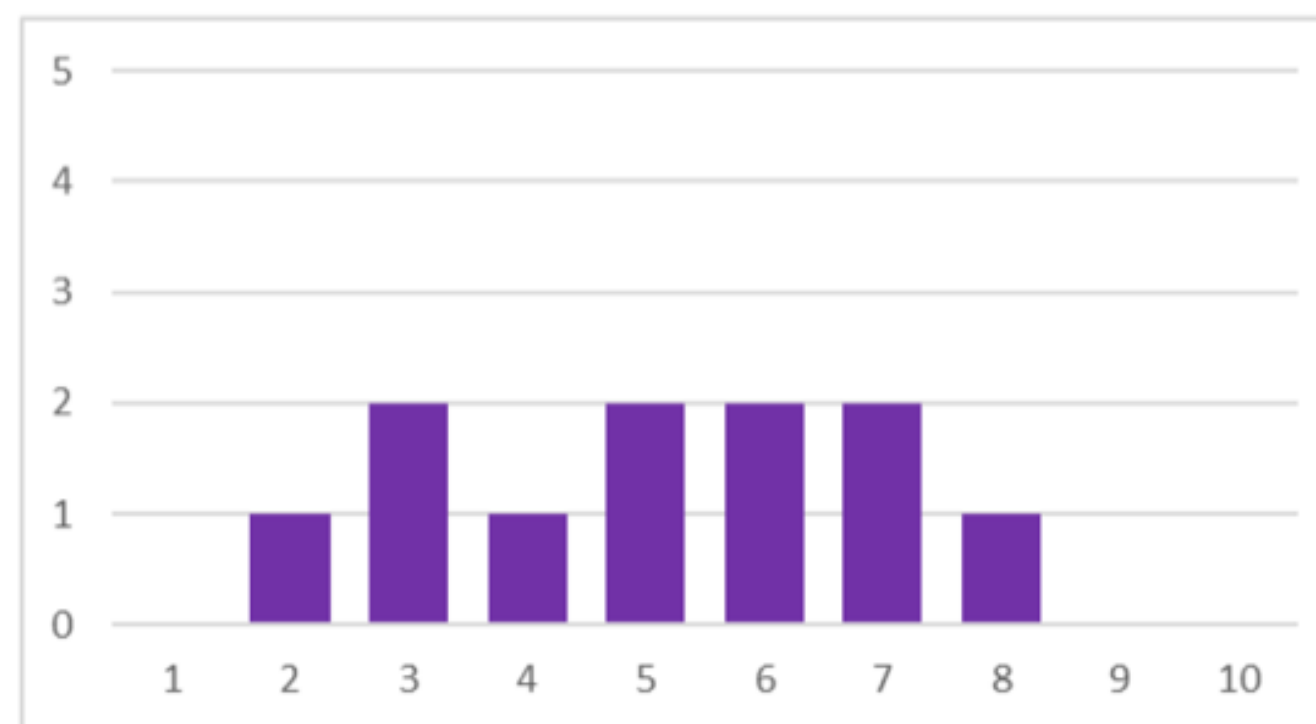


Figure 12. Frequency histogram of audio group

How likely are you to study in the future?



# Quality Control Is a Challenge

- Some students have specific requests

Student: "I do not want to read about accidents and deaths!"

- Conflicting goals: recency, difficulty, interest

- Some articles are paywalled

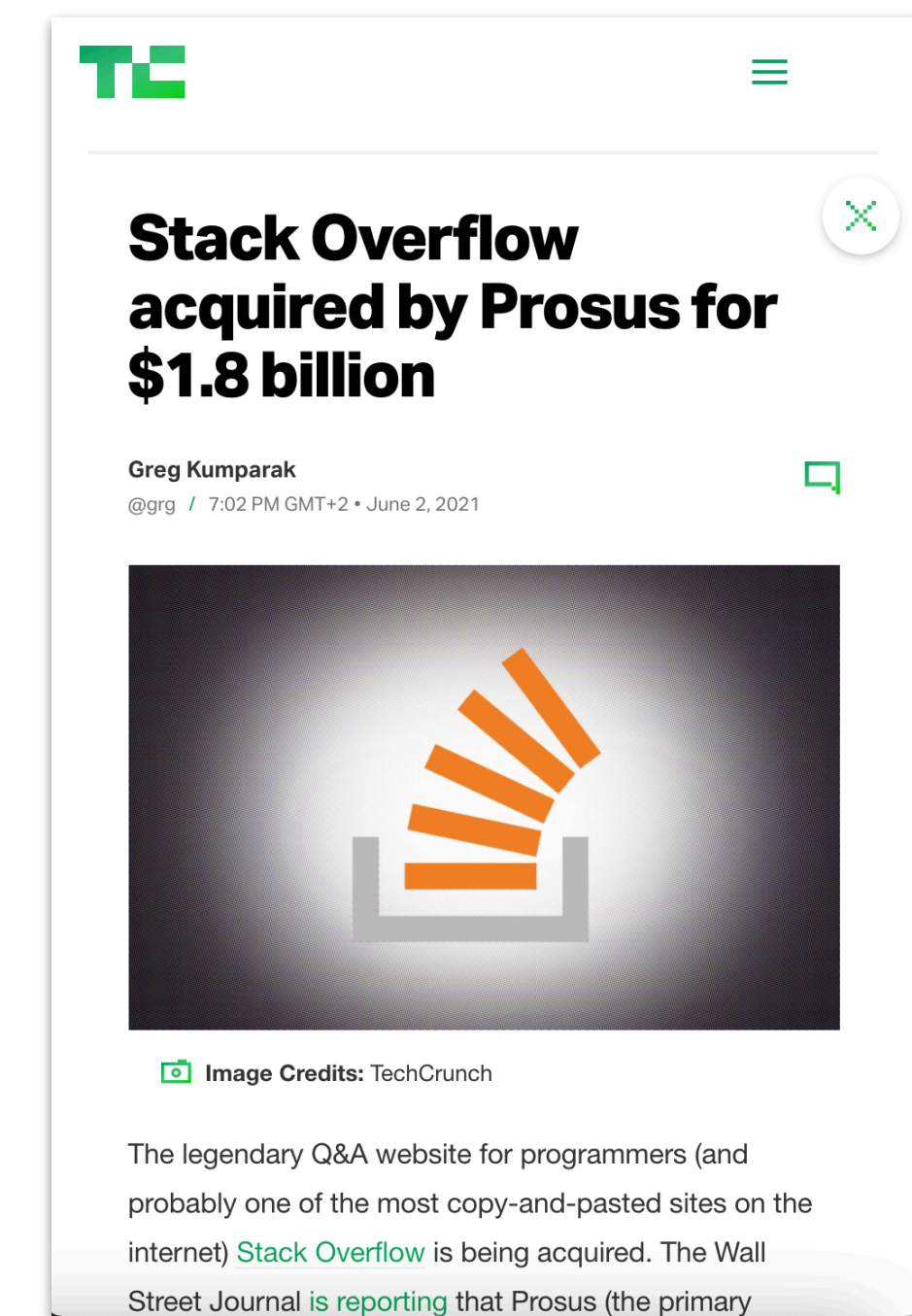
Eating at McDonalds (Textbook) vs local when traveling...

## Quality Control Forward

Semi-automatic  
(Apple News)



Crowdsourcing  
(StackOverflow, Reddit)

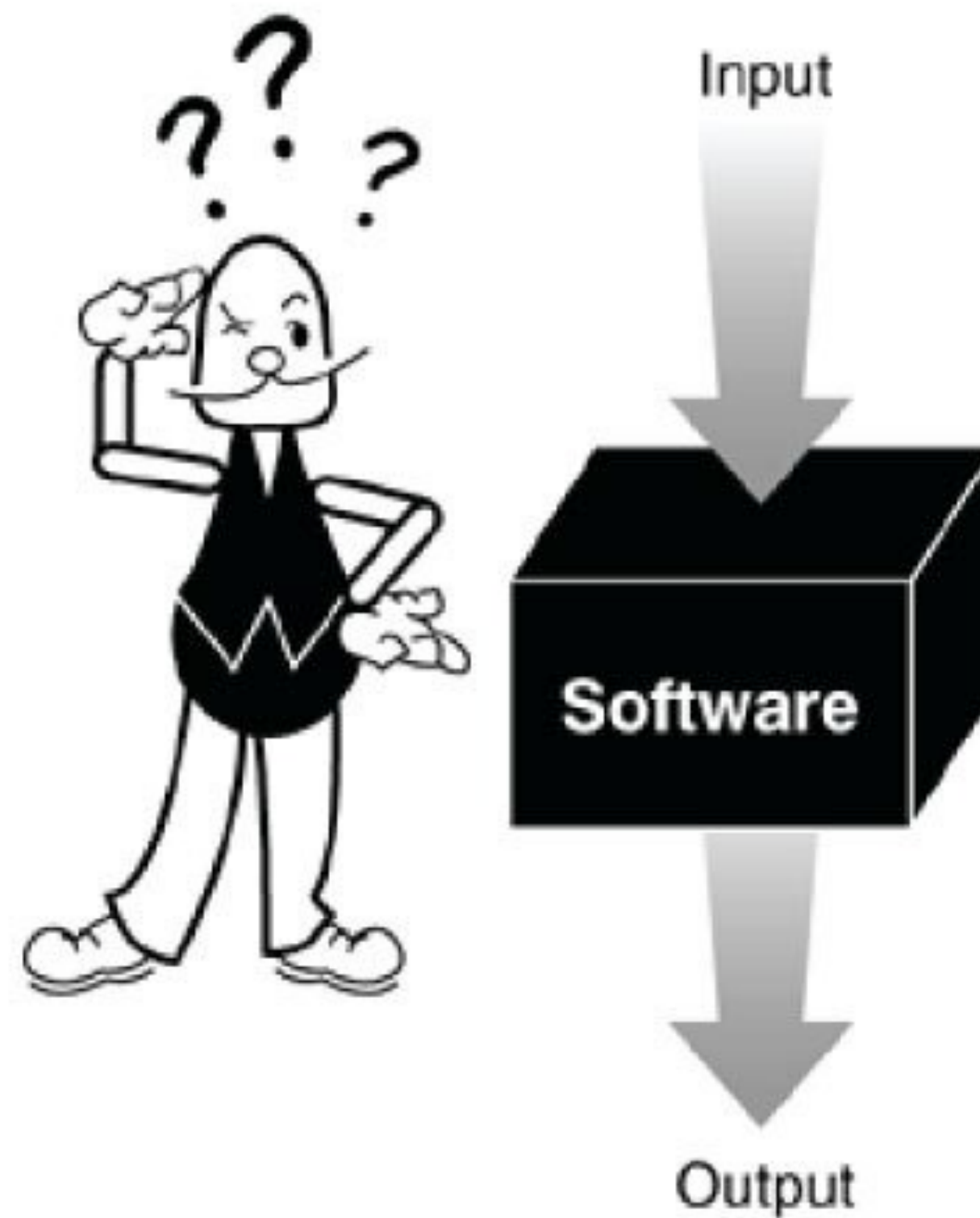


# Teachers want visibility into (some) algorithms

“Why are there some words that the students don’t see in exercises?”

“What is the definition of learned words?”

“How is reading time computed?”



NB: teachers don’t care about the algorithm for assigning categories...

# Cleaning up news pages of irrelevant graphics is not easy

- Mozilla Readability
- Mid-text images
- Mid-text textual ads
- Read also...

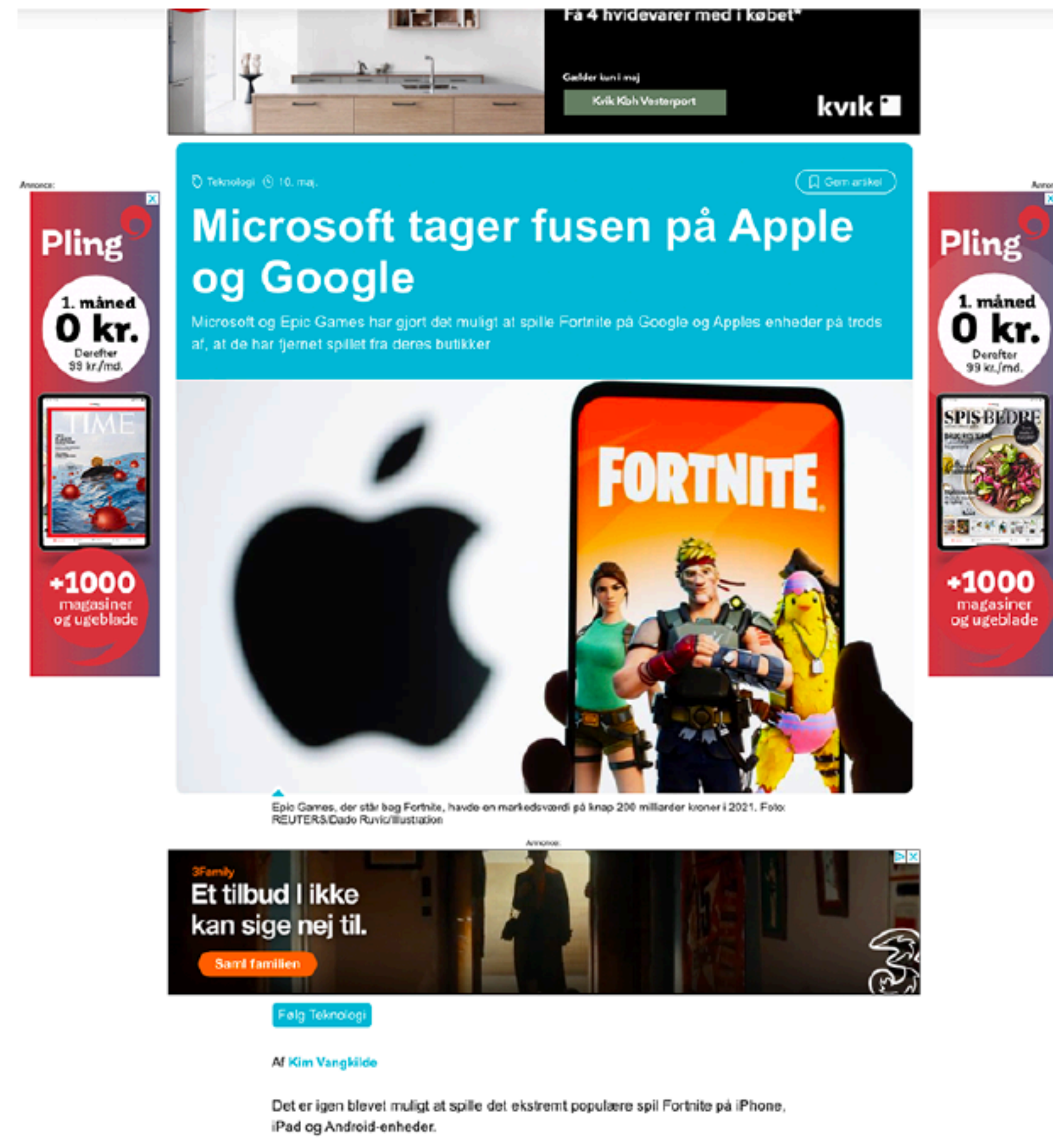


Figure 9: On the left is the original article on the website. On the right is the same article in the reader view

# But in the end... just a tool in the toolbox

- The user that keeps reading two years later :)  
The learner that does audio every day almost!
- Some teachers succeed to use the tool in the classroom for multiple years now
- Student showing the teacher “*how many words they translated*” >-)



# Future Directions

## Short Term: Improve

- **Difficulty estimation (with Michalis & Yash)**
- Knowledge modeling
- Progress feedback

## Mid Term: New Features

- Video, Audio, Social Media
- Social aspects
- Mobile applications
- Commercialization (UCreate)

## Long Term: Fine-grained Personalization

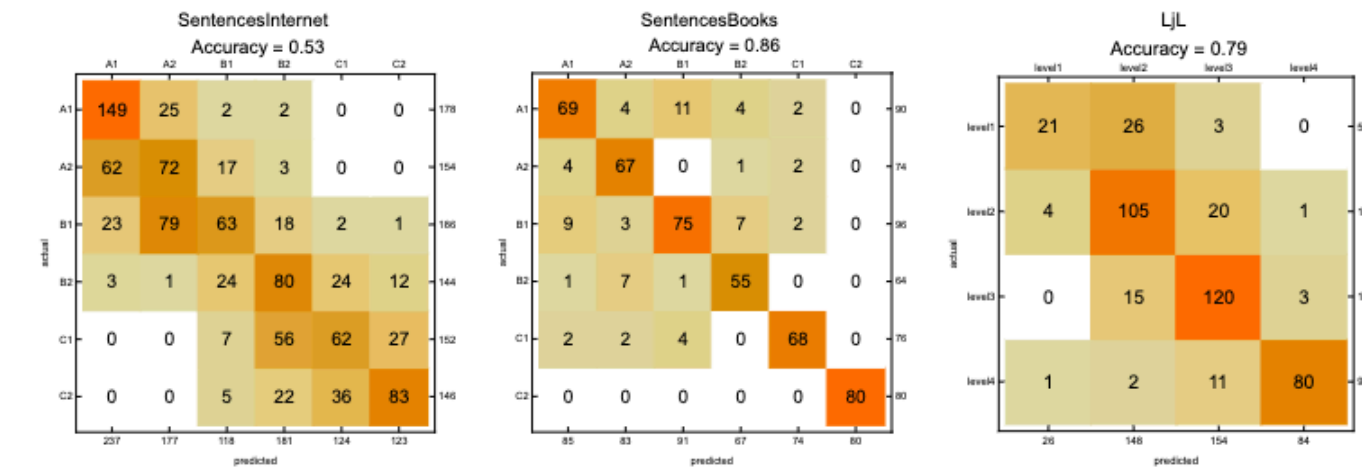


Table 2: The classification accuracy and confusion matrix for the difficulty estimation technique based on the Bert embeddings for the three datasets used in our experiments.

Dataset	Our approach	GFI	ARI	FKGL
LjL	<b>11,354</b>	17,338	17,749	19,634
sentencesInternet	<b>67,288</b>	83,770	92,140	85,306
sentencesBooks	<b>9,875</b>	30,836	32,993	31,475

Table 3: Number of pairwise mismatches for our difficulty estimation technique compared to predominant readability metrics. Numbers in bold indicate fewer mismatches, and thus better estimation of difficulty.

annotated with a difficulty level by a Professor of French. All sentences in that book were then given that label. This process involved an OCR pipeline which could lead to faulty detection of characters, so only those sentences without any errors were retained. (labels: A1,A2,B1,B2,C1,C2).

The characteristics of these datasets are provided in Table 1. To train and evaluate our model, we use an 80/20 train-test split.

### 4.1 Accuracy

We compare the accuracy of our difficulty estimation approach to traditional readability based metrics, such as the GFI (Gunning Fog Index), FKGL (Flesch Kincaid grade level), ARI (Automated Readability Index). These techniques are inherently regression techniques and output a floating point value of the text difficulty, therefore we cannot make a direct comparison, because our difficulty estimator predicts discrete labels.

To make a meaningful quality comparison, we devise the following experiment which draws on **pairwise-comparisons**. Then, for each text with a specific label, we compare its difficulty to all the other objects of the other labels and record if this comparison was correct. For example, if a text with label A1 received an ARI score of 15 and a different sentence with label B2 received an ARI

score of 14.6 (i.e., indicated as easier), then this comparison is considered incorrect and recorded accordingly. We call this incorrect comparison a **mismatch**. With our approach we can also do the same by comparing the labels. If the confusion matrix of a classifier is already computed, then the pairwise mismatches can be evaluated as shown in Figure 5.

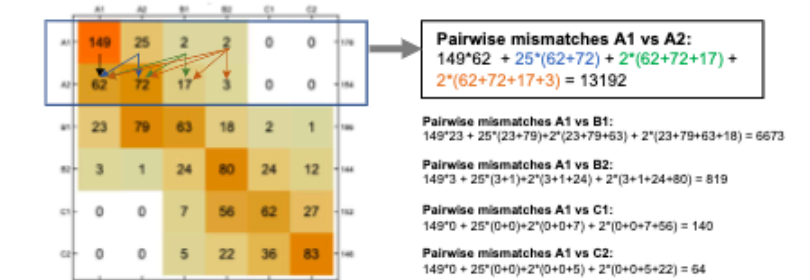


Figure 5: Evaluating the pairwise mismatches for a given confusion matrix of a classifier (sentencesInternet).

The results of this analysis are shown in Table 3. Higher numbers indicate more mismatches. Therefore, we see that compared to older or traditional methods of difficulty estimation, the approaches based on modern embeddings offer a significantly lower error rate.

“Such a system is **critical for language education in schools**, since the possibility of choosing their topics of interest is motivating”  
(French Language Teacher)

*“My vocabulary truly is improving but you have to use it more than a few times. **Easy to use and with nice topics**”.* (Student)



**Try out the system**

Invite Code: **ec25**

<https://zeeguu.org>



**Source code, papers, datasets**

<https://mircealungu.com/projects/zeeguu>



**Get In Contact**

[mircealungu.mlun@itu.dk](mailto:mircealungu.mlun@itu.dk)

[mlun@itu.dk](mailto:mlun@itu.dk)