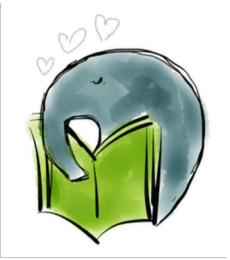


Zeeguu - Personalized Learning Paths Through the Internet

Mircea Lungu, IT University of Copenhagen
Sake Jager, University of Groningen

EUROCALL 2025 – Milano
Università Cattolica | 27 August 2025

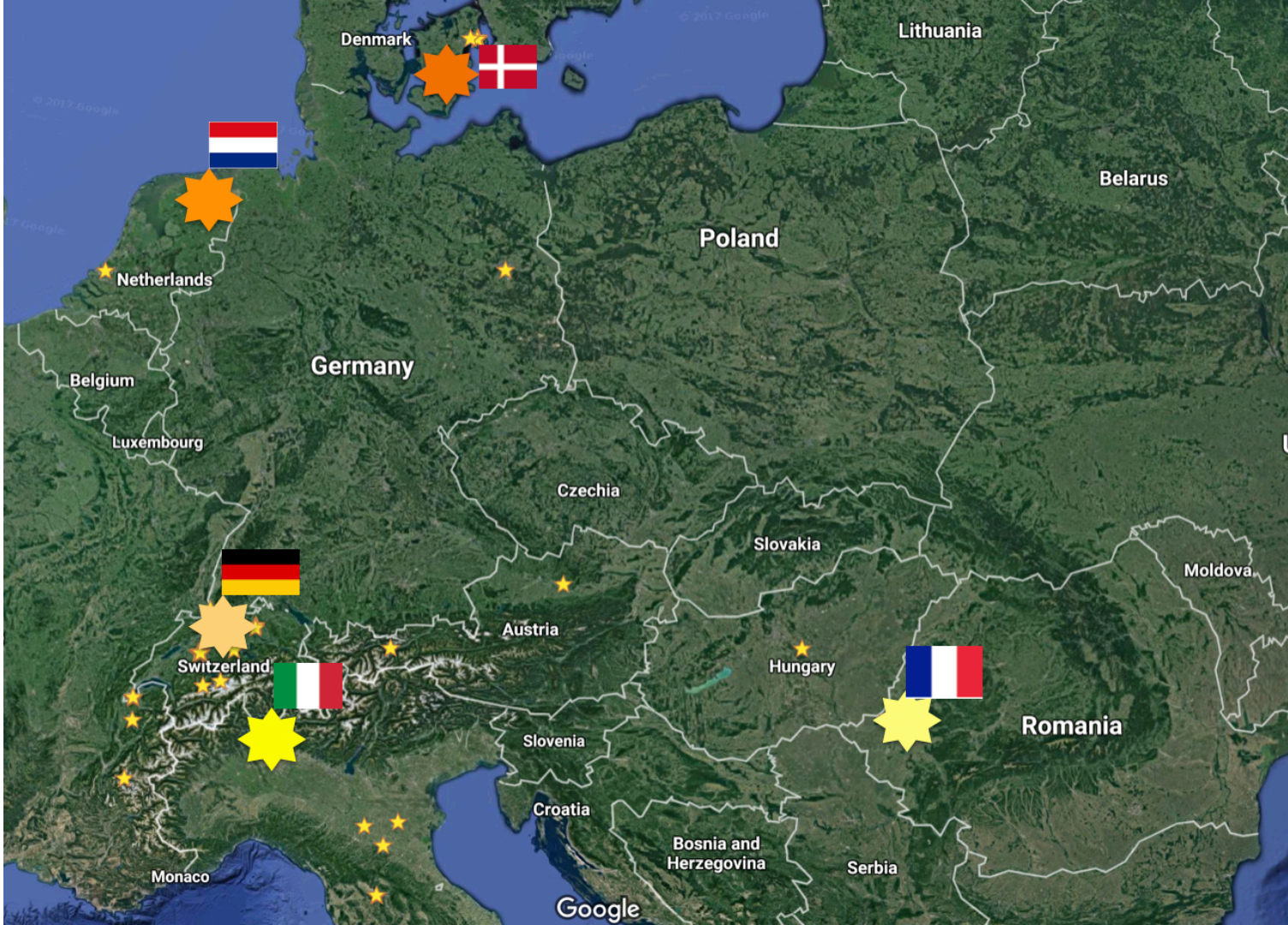
About us

Mircea Lungu

- Associate professor at IT University in Copenhagen
- Main specialization - software engineering, CALL is a secondary line of research
- Working on Zeeguu since many years with many students - each developing a small aspect
- Started it because wanted to learn German
- Using it to learn Danish

Sake Jager

- Project manager Teaching & Learning Innovation, Assistant Professor Applied Linguistics, University of Groningen
- CALL Course
- Projects with Zeeguu from 2016
- Use by colleagues and students
- CALL project 2025: evaluate autonomous use in context of L2 Dutch course



Always
Learning
A New
Language

Vision: Hyper-Personalizing the Internet for Learners

Using personalization *for good*

1. Personalized article recommendations based on **difficulty and topic**
2. Minimal-effort **interactive translations** to compensate for the imperfection of recommendations
3. **Optimal vocabulary consolidation** with exercises based on past interactions in readings

as opposed to…

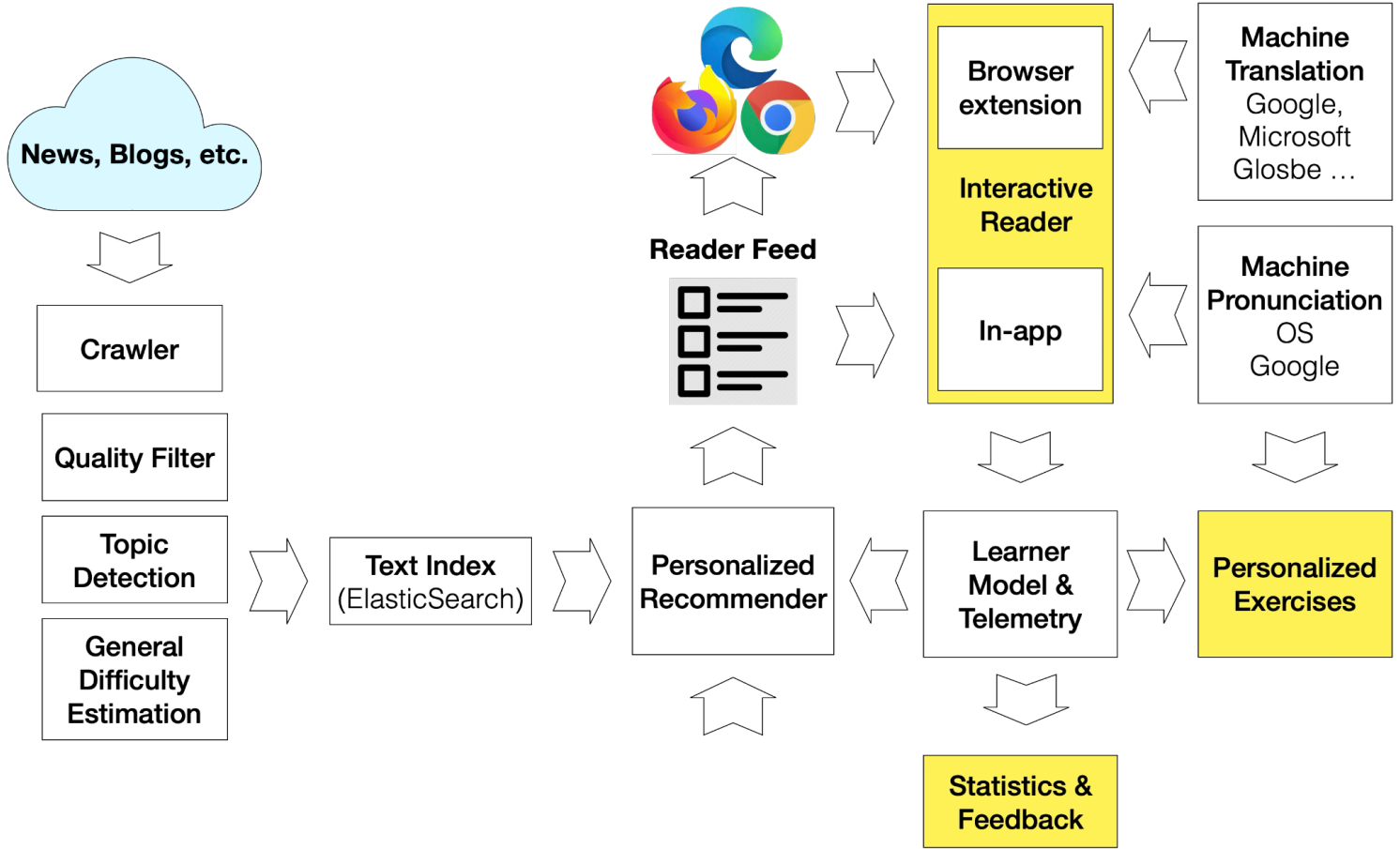
Part 1: Interactive Presentation

Reading with Zeeguu

- Demo by Sake on:
  - Settings
  - Selecting and reading articles
  - Choosing words for learning
  - Additional features
- Presentation by Mircea of:
  - Article recommendation
  - Difficulty estimation
  - Topic assignment
  - Translation
  - Simplification

How do we recommend articles?

we seed this
with sources we trust



Difficulty Estimation - Traditional with FK

- Calculates three metrics
  - Words per sentence,
  - Syllables per word*
  - Total sentences
- Uses language-specific formulas
  - e.g., English: 206.835 - 1.015×WPS - 84.6×SPW)
- Converts to grade between 0-100 (Inverted FK score)
- Usability feedback
  - Users don’t care about 67 vs. 68
  - Switched to CEFR levels – A1 (>83) to C2 (<15)

* Using *pyphen* library for accurate hyphenation

Difficulty Estimation - LLM-based

- Since two months we are experimenting with an LLM for estimating difficulty
- We still need to evaluate it

This Danish text is at **CEFR level C1 (Advanced)**.

Here's why this text qualifies as C1 level:

Vocabulary complexity:

- Highly specialized scientific terminology (kosmologi, galakser, lysår, udvidelse, tomrum)
- Abstract concepts requiring advanced comprehension
- Technical phrases like "kosmologiske model" and "Hubble Tension"

Grammatical structures:

- Complex sentence structures with multiple subordinate clauses
- Passive constructions and advanced verb forms
- Sophisticated conjunctions and linking words

Text characteristics:

- Academic/scientific discourse with formal register
- Dense information requiring sustained attention
- Abstract reasoning about cosmological theories
- References to scientific literature and research

Cognitive demands:

- Understanding of complex cause-and-effect relationships

Topic Assignment? Three Types of Topics

1. **Website-Specific** Topics
    a. Predefined article feeds: The Onion, Lercio, Der Postillon -> “satire”
2. **URL Keyword-Based** Topics
    a. Some urls we know for sure the topic (e.g. **/politics/article** → "politics")
3. **Content-based Inference**
    a. Uses semantic similarity via Elasticsearch
    b. Finds k=9 similar articles using text embeddings
    c. - Collects topics from these similar articles
    d. - If the most common topic appears in ≥50% of neighbors, assigns it

Catering to more specific interests

Searches

- Keyword-based
- Language specific (e.g. København for Danish)
- **Saved** searches are automatically included in recommendations

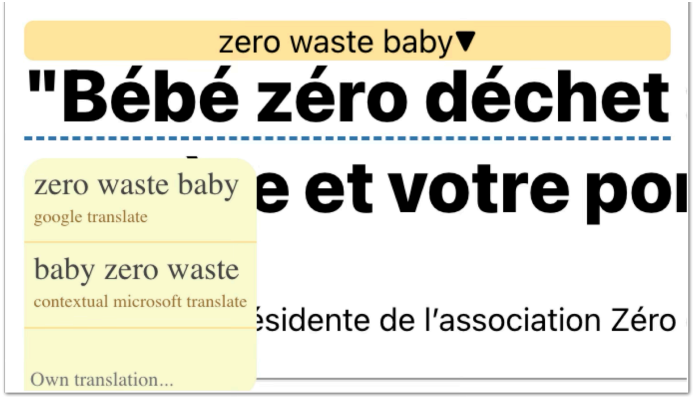
Exclusion filters

- Keyword based
- Language specific (e.g. død for Danish)
- Disturbing news

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

Translations

- Machine translation from multiple APIs
  - Learner can choose alternatives
  - Learner can upload own
  - Needs to understand enough of the context to choose the right one
- (Word, translation, context) saved in DB
  - Context is tokenized with Stanza (sentence end for context extraction)
  - Translation saved together with token index to DB (in case word occurs twice in text)



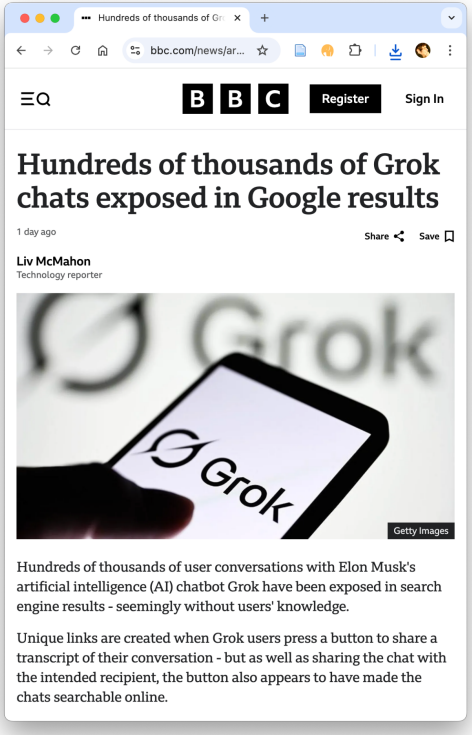
Approx. 1/7 translations => dropdown

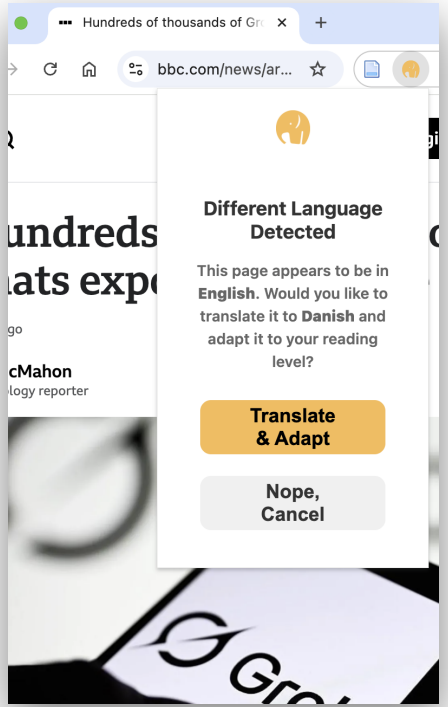
The teacher who liked the imperfection

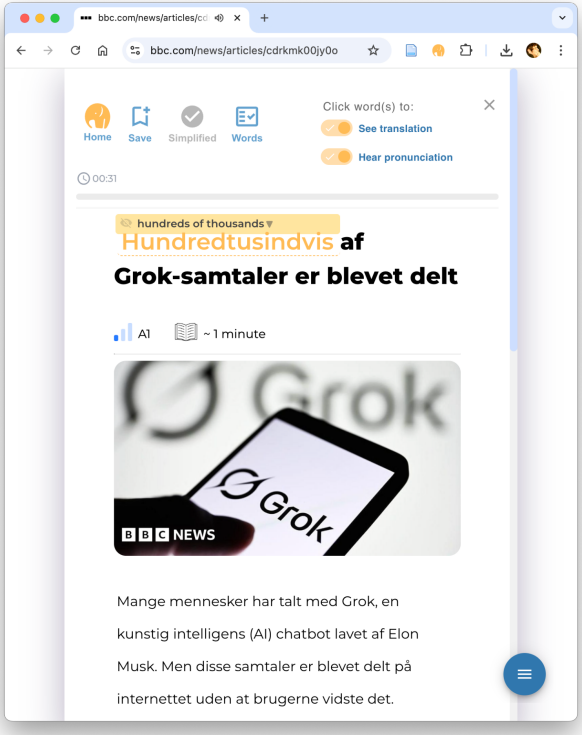
Simplification

- Problem: **very few articles on the net are A1 or A2**
- Idea: what if we have the LLM automatically simplify existing articles?
  - Same LLM as difficulty estimation
  - For every article we save the simplified versions in the DB
  - When the article would be recommended based on the interests, we swap it in the recommendation feed instead of the original
- Limitations
  - Not evaluated yet
  - Not perfect (machine job)
  - A1, A2 – more related to personal life - we aim to do the simplification for any topic!

Translation + Simplification = New Feature







Towards a lens over the whole Internet?

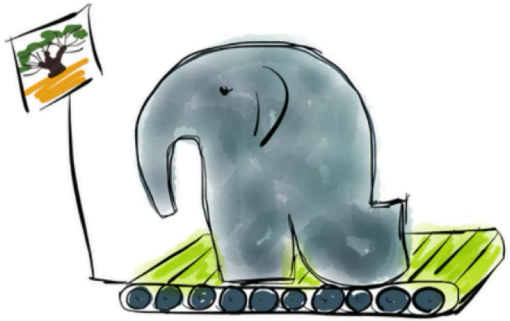
Vocabulary Practice

Demo by Sake on:

- Exercises: Matching, gap-fill, select in context
- Audio: Words in spoken contexts
- Other features: Words, History, Statistics

Presentation by Mircea of:

- Exercise types and levels
- Scheduling
- Generation of audio lessons
- Alternative examples



10 Exercise types in 4 levels

Based on two dimensions

i. **Cognitive Demand**: Recognition (selecting from options) vs. Recall (producing from memory)
ii. **Skill Type**: Receptive (understanding L2) vs. Productive (generating L2)

Level 1 – Recognition/Receptive (Easiest)

Understanding L2 with multiple choice support:
- **MultipleChoiceL2toL1** – See L2 word, select L1 translation
- **MultipleChoiceAudio** – Hear L2, select from options
- **Match** – Match L2 words to L1 translations
- **ClickWordInContext** – Click on target word in L2 text
- **FindWordInContext** – Identify target word in L2 text

Level 2 – Recognition/Receptive (Context-based)

Understanding L2 in context with options:
- **MultipleChoice** – Read L2 context, select fitting L2 word
- **MultipleChoiceContext** – See L2 word, select fitting L2 context

Level 3 – Recall/Receptive

Understanding L2 and producing L1 (no options):
- **TranslateL2toL1** – See L2, type L1 translation
- **TranslateWhatYouHear** – Hear L2, type L1 translation

Level 4 – Recall/Productive (Hardest)

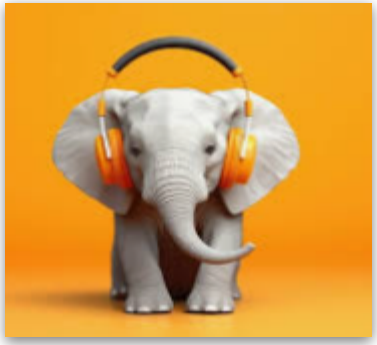
Generating L2 from memory:
- **SpellWhatYouHear** – Hear L2, type L2 spelling
- **FindWordInContextCloze** – Fill blank with L2 word

Exercise Scheduling

- Spaced repetition is more challenging with multiple exercise types
  - You progress the word every time it's correct
  - A word has to go through four levels before is learned
  - Each level has two different exercises (1 day and 2 days cooling interval)
  - With cooling intervals: and then you move
  - The fastest to learn a word: 8 times correctly, about two weeks
- Maximum number of words the user can have In learning at a given point
- As words are learned, **new words are introduced based on**
  - Quality checks (expressions not more than 3 words)
  - Word rank (lower ranks have priority)

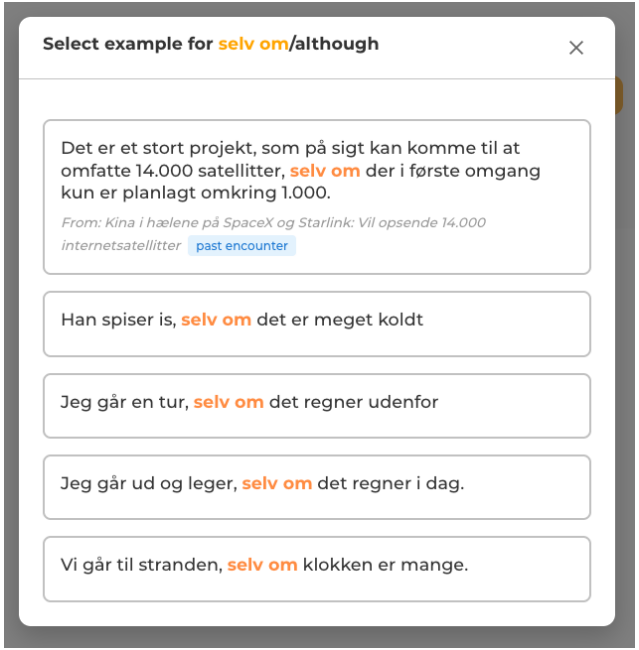
Audio Lessons - Most Recent Experiment

- Also use the LLM
- Every day automatically chose three main words to focus on
- Lesson structure
  - Short dialogue
  - "*Repeat after me*" kind of challenges
    - Sentences on the dialogue
    - A few more sentences
- Based on a traditional approach
  - No speech recognition - too difficult with non-natives
  - Comprehensive **input** is our goal
- Evaluation: TBD!
  - Two users really *"hooked"* - Mircea and a Dutch learner



Alternative Examples In Exercises

- The original context is not always the best
  - Some sentences do not make sense on their own
  - Some sentences are too long and make exercises unpleasant
- Solution
  - Create a **DB of alternative examples** for a given word/translation pair
  - **Allow the learner to change** the example that is used in exercises
  - Preliminary, subjective, & ongoing evaluation
    - Good because otherwise one memorizes the solution based on context
    - Useful to see alternative uses



Part 2: Participants Try the System Individually or in Groups

Go to: https://zeeguu.org/
Click on **Start Learning**
Use invite code: **Meetup25**

Part 3: Participant-Led Discussion

Your experiences and interest

- Was there something that was not clear
- What did you think about the app?
- How would you use this in your own context?
  - as learners
  - as teachers
- Would you be interested conducting research on Zeeguu use?
  - Topics?
  - Specific features of interest?

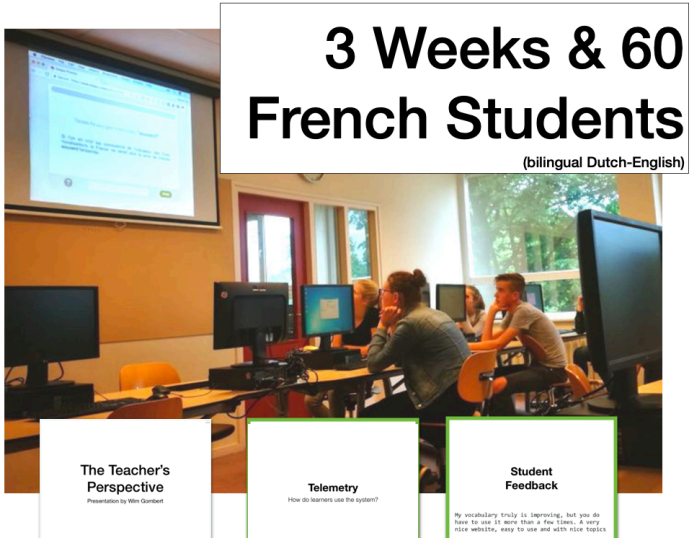
Further Discussion Points

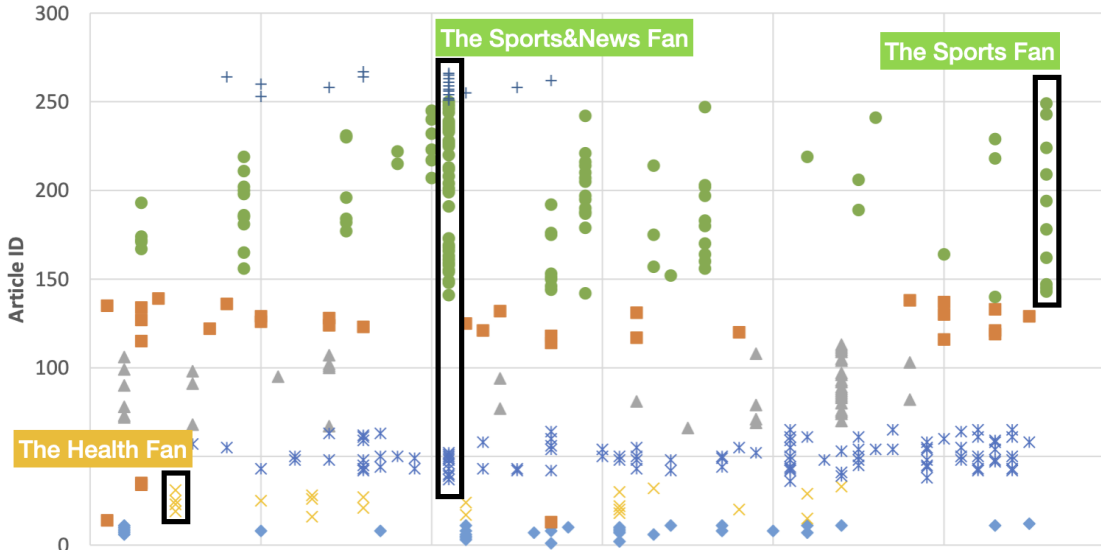
Experiences

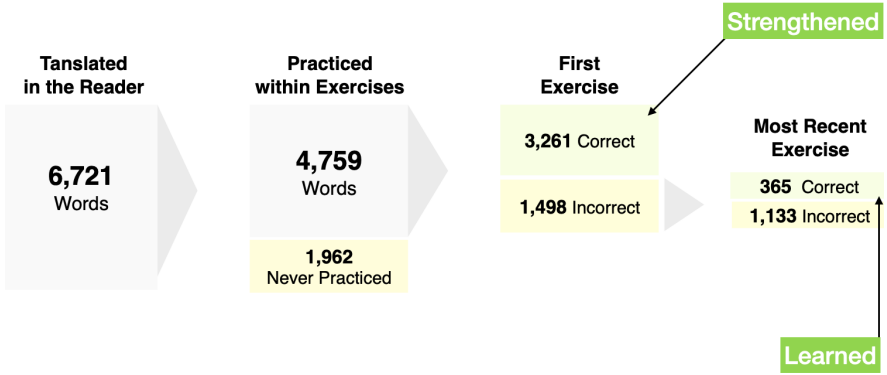
Our own experience

- Mircea
  - Dutch highschool students reading French
    - Multiple highschools
    - Using the platform for many years now
    - Task is - free reading at home
    - Reward is - 0.5 point for 1/2h
  - RUG Dutch learners
    - Using the platform for multiple years
    - Retention is low - one in 50 learners stays with it for the long term
  - Various studies with master students

French-learning Dutch highschool students







I would like to **avoid articles** with information about **accidents** with **human casualties**

Several small studies with learners

- Participants appreciate the browser extension
- Participants consider that audio exercises make the system more fun
- Inconclusive study on the impact of progress visualization across learning levels for a word
- Study on augmenting video the way we do articles - participants are positive

https://github.com/zeeguu/studies

Possible collaborations

Learner Studies

- Impact of simplification - systematic study
- How to show meaningful progress to the learner
- Impact of audio lessons - I believe it is going to be big!

Prototyping

- If you have ideas of small, vertically integrated features that you would want evaluated, we can discuss

Data Analytics (thousands of learners, anonymized info)

- Do people actually learn?
- Are our scheduling algorithms optimal?

Challenges

Future Work

Open Questions

- How to provide meaningful progress feedback?
- How do we evaluate the simplification feature?
- How do we evaluate with LLMs - they are ever changing?
- Ethics of presenting simplified articles directly in-app?
- Quality of simplifications?
- How can we collaborate with others?
- How do we compensate for the imperfection of the ML tools?

Many People Involved

Tiago Ribeiro
Development and ML Research

Iga Anna Waclawska
Onboarding Experience Improvement

Merle Schön
Exercises and Learning Algorithms

Darinka Maldonado
Browser Extension Usability Improvements

Beatrice Verzier
Exercise Improvements

Bjørn Hjorth Westh
Danish Localization

Gustav Norgaard Petersen
Audio Exercises

Jonas Krohn
Audio Exercises

Maria Timis
User Dashboard

Emma Lerche Teglbrænder
Browser Extension

Frida Helene Beck-Larsen
Browser Extension

Sara Tolstrup Husum
Teacher Dashboard Modernization

Eleonora Kurilchik
Teacher Dashboard Modernization

Maja Albin Groes Ludvigsen
Teacher Dashboard Modernization

Katrine Iversen
UI Modernization

Konstantina Argyropoulou
UI Modernization

Kristin Kallevik
UI Modernization

Feiko Ritsema
Topic Browser

Lars Holdijk
API Design

Dan Chirtoaca
Universal Multilingual Reader

Alexander Lukjanenkovs
Learner Feedback

Anca Lungu
Linguistics Advice

Linus Schwab
Android RSS Feed Reader

Jeroen van Engen
Education Feedback

Mads Kristian Brodt Nielsen
Teacher Dashboard

Sybren van Vliet
Forgiving Exercise Feedback

Martin Avagyan
Language Exercises

Luc van der Brand
Universal Multilingual Reader

Jorrit Oosterhof
iOS News Reader

Karan Sethi
Knowledge Estimator

Pascal Giehl
Android Development

Simon Marti
Initial API and Chrome Extension

Wim Gombert
Education Feedback

Alin Balutoiu
Performance and Scalability

Jonathan Sparvath
Teacher Dashboard

Joël Grondman
Personalized Difficulty Estimator

Carlos Paz
Knowledge Dashboard

Johan De Jager
Translation Multiplexer

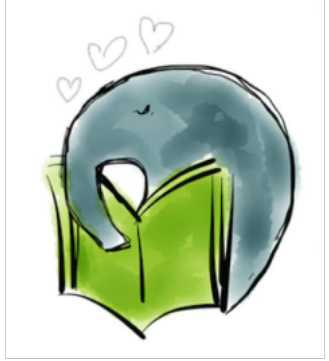
Ada Lungu
UX and Web Design

Niels Haan
Smartwatch Development

Vlad Turbureanu
Language Exercises

"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.com
mlun@itu.dk

Backup Slides

From translations to meanings

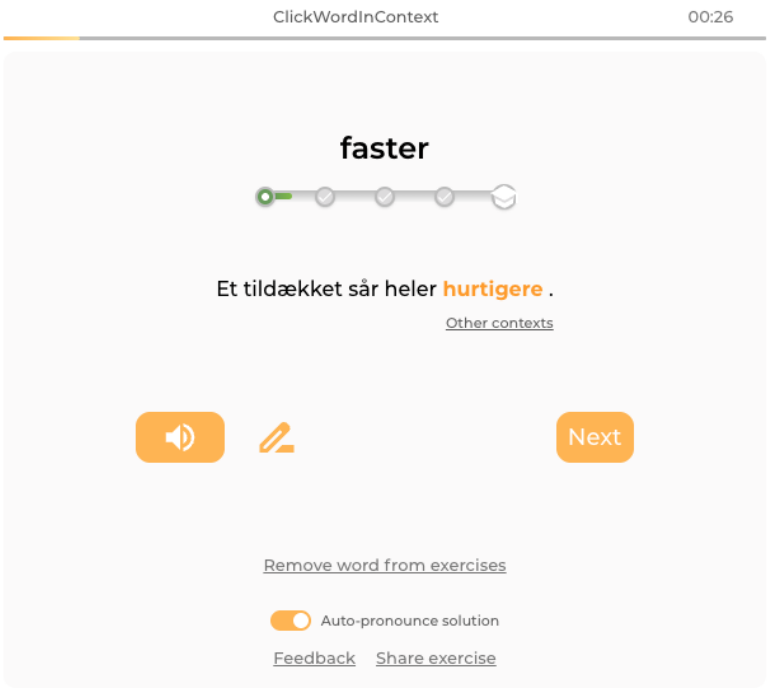
- Initially every translation would be scheduled for learning
- But then, if a learner translated the same word twice, then it would be scheduled twice
- To avoid this, we model in our DB distinctly the **translation** and the **meaning** it points to
  - Challenge: mapping syntactic variation on the same meaning
    - *Oh, this is quite an august entourage.*
    - *August entourage is what we have here, I'm telling you.*
  - Challenge: disambiguate between meanings
    - *Let's invite August to tag along!*
    - *I proposed we do our vacation in August.*

Richer semantic modelling?

But if we know that the learner already knows the word **hurtig** (i.e., fast) then maybe we don't even have to schedule the word **hurtigere** (faster)?

Then again… learner **did** click on it to get a translation… so maybe they should learn it.

Also, if it's too easy they can **remove it from exercises**



In Learning (48) | **Learned** (149)

Congratulations! You have learned **149** words so far.

A word is learned when done correctly in exercises in all the four levels. Or when it is marked as *"too easy"* by you after you have done several exercises with it.

Vocabulary Progress?

August 2025 (36 words, median rank: 1311)

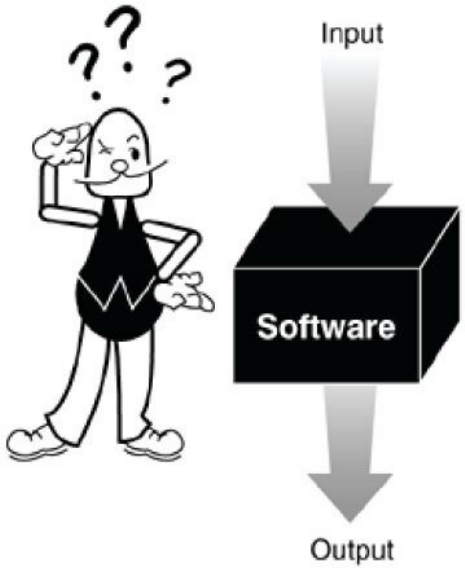
utrolig 1639 - incredibly
ved 32 - at
løfte 1682 - promise
desuden 1671 - also
undersøge 1619 - investigate
overstået 1251 - ended
virker 351 - seems
gjorde 128 - did
som 36 - like
ødelægger 1384 - ruins
væk 132 - away
målet 1550 - the target
forsvundet 1424 - disappeared
gemmer 1373 - hide
hemmelighed 1344 - secret
næsten 362 - almost
behøver 250 - need
intet 173 - nothing
bud 1662 - bid
angrebet 1634 - attacked
skilt 1614 - divorced
sager 1463 - things
skød 728 - shoots
ligner 409 - look like
Disse 301 - These
ældre 1293 - older
har tit 1224 - have often
vender 960 - turn over
halvdelen 1461 - half
overleve 1331 - survive
forestille 1328 - imagine
dårlige 1247 - bad
overhovedet 806 - even
forventer 1431 - expects
støtte 1409 - support
kunne 70 - could

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…