





On using AI methods to carry out SLA research in real-life school contexts

Detmar Meurers

University of Tübingen & Leibniz-Institut für Wissensmedien (IWM)

based on collaborative research: http://icall-research.de

Language and Speech Laboratory, University of the Basque Country Micaela Portilla Research Centre, Vitoria-Gasteiz, March 19, 2024

Overview

- Second Language Acquisition Research:
 - insights from multiple strands of SLA research and some desiderata
 - instructed SLA: from "nice to have" transfer of SLA research results to crucial large-scale data source needed for research on ecologically-valid language learning!
- Digital learning contexts:
 - increasingly widely used in real-life teaching and learning
 - adding AI methods to make digital tools more effective in practice and useful for SLA research
- Scaling up SLA intervention studies using digital means
 - Web-based studies
 - Randomized Controlled Trials in schools



- long history (Corder 1967; Selinker 1972), EuroSLA since 1989/1991, SLA recognized as a distinct discipline for over 20 years (Larsen-Freeman 2000)
- identified a wide range of linguistic, cognitive, social factors influencing SLA, including:
 - Input Hypothesis, i+1 Hypothesis (Krashen 1985)
 - Noticing Hypothesis (Schmidt 1990)
 - Teachability Hypothesis (Pienemann 1984)
 - Zone of Proximal Development (Aljaafreh & Lantolf 1994)
 - Output Hypothesis (Swain 2005)
 - Interaction Hypothesis (Long 1980a; Pica & Doughty 1985)
 - Cognition Hypothesis (Robinson 2003) in Task-Based Learning
 - ...

EBERHARD KARLS

Second Language Acquisition Research: Some challenges

- limited operationalization and empirical breadth of some key notions:
 - i+1, developmental sequences, zone of proximal development, ...
- in real-life learning, linguistic, cognitive, and social factors are simultaneously involved
 - Long (1980a): "SLA is a **multidimensional** phenomenon, involving (at least) biological, neurological, social, psychological, affective, and personal variables plus [...] the effect of instruction."
 - DeKeyser (2012): important to study **interactions** between individual differences, treatments, structures
 - But the reliable identification of interaction effects require a lot of data!



Second Language Acquisition Research: Obtaining rich data at scale

- How can we obtain more data on the learning process and product given a range of factors?
- 34 million children in secondary schools in Europe are learning English as L2 and 20 million children are learning more than one foreign language (Eurostat 2020/21)
 - relatively understudied population in SLA research
 - legal and practical issues make it difficult to stage interventions and collect data at scale
- digital learning contexts offer new opportunities for collecting data

Digital Learning Contexts

EBERHARD KARLS

TÜBINGEN

Schools going digital is a common topic in public discourse, e.g.:

Computers in all schools, all students at the computers – this is the program the ministers of education want to realize quickly. [Spiegel No. 47 / 18.11.1984]

- More recent game changers:
 - widespread use of mobile digital devices (smartphones) by children
 - Corona pandemic required systematic use of digital channels
 - Al methods support development of tools linked to SLA concepts



Digital learning contexts: from digital media to adaptive, interactive tools

- Traditionally, computers used in education encode and exchange digital information to
 - increase motivation through authentic multimedia material
 - support **communication** with native speakers (using text, audio, video)
 - **mechanically practice** vocabulary and grammar (without modeling the learner or providing feedback)
- To obtain interactive, adaptive tools, one needs to add Al methods that can
 - analyze and generate **language** (using computational linguistic methods)
 - model **learners** and their development towards the curricular **learning goals**
 - analyze and adaptively select or generate **learning activities**
 - → Intelligent Computer-Assisted Language Learning (ICALL)
- ⇒ Explore ICALL tools linked to SLA concepts to
 - effectively support learners and teachers
 - research SLA hypotheses in real-life contexts

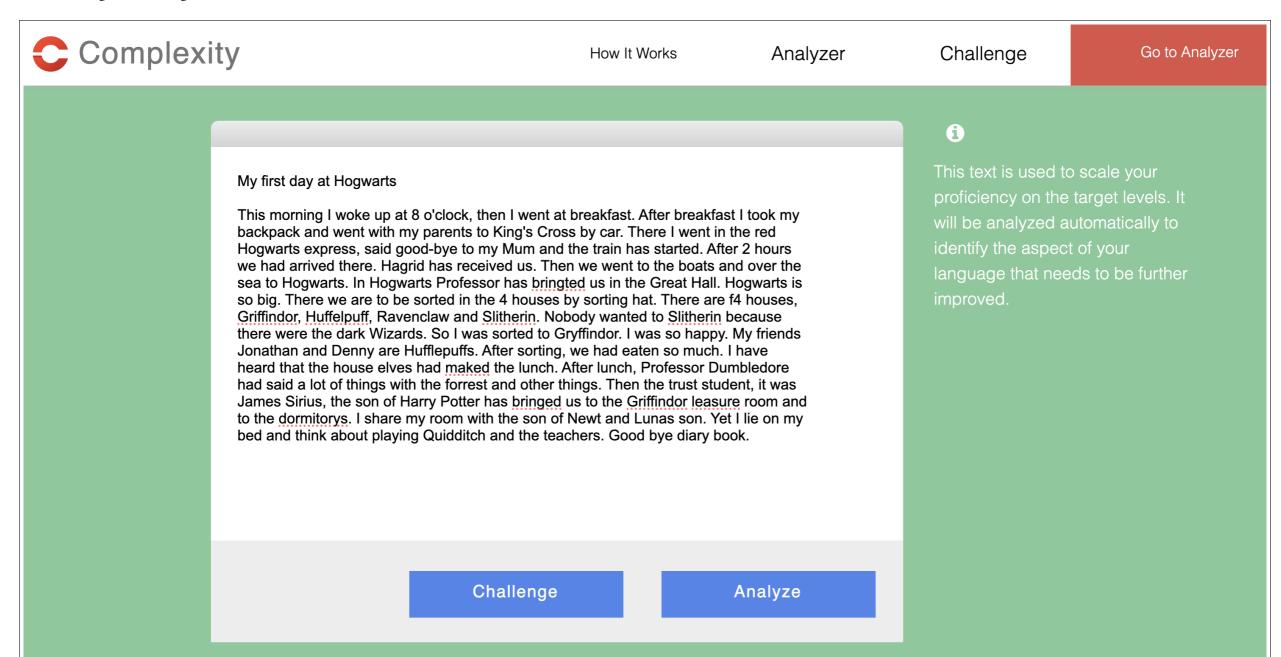
Key Factors: Input and i+1

- Input Hypothesis (Krashen 1985)
 - access to comprehensible input at i+1 is essential
 - input should be "one step more complex" than the interlanguage of the learner
- \Rightarrow ICALL tools can help us identify linguistic complexity to search for such i+1 input:
 - SyB System (http://complexity.schule, Chen & Meurers 2019)
 - FLAIR System (http://flair.schule, Chinkina & Meurers 2016)

7/43

EBERHARD KARLS

SyB: Syntactic Benchmark (http://complexity.schule, Chen & Meurers 2017)



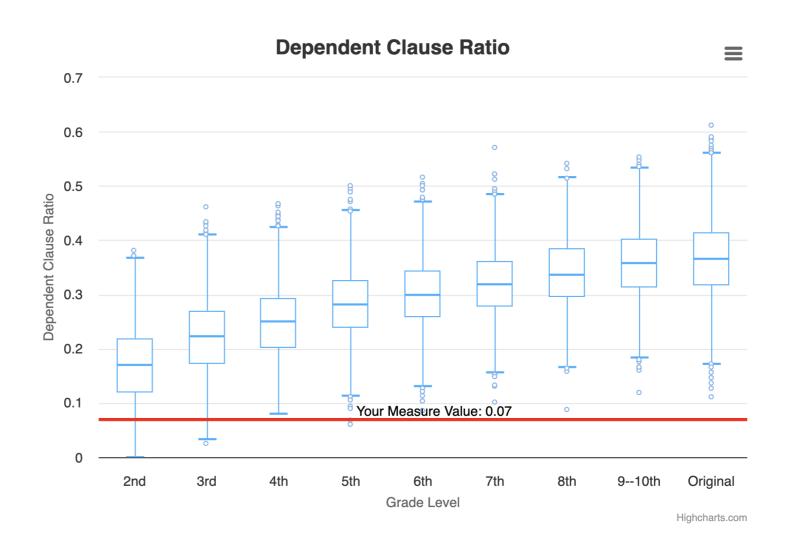
How It Works



C Complexity

Dependent Clause Ratio

Challenge



Syntactic Measure:

Mean Length of Clause

Challenge



All Grade Levels

Challenge Level





Middle-class Latinos "step it up a notch" as communities thrive

When Rick Rodriguez Sr. moved into a three-bedroom home in Downey, California, in the mid-1980s, a neighbor interrupted his yard work with a question: ?You?re not going to be one of those guys that parks in the front yard and lets chickens run around in the front yard, are you?? Rodriguez, 54, one o...

Grade Level: 8th

Target Measure: 10.19



11 of 12 defendants in Atlanta schools case found guilty

ATLANTA? After considering nearly five months of testimony, a Fulton County jury convicted all but one of the dozen former Atlanta teachers and administrators on trial in the largest test-cheating scandal in the nation. On Wednesday afternoon, 10 former teachers and school administrators were handc...

Grade Level: Original

Target Measure: 9.27

Target Measure: 10.05



Rude political candidates conflict with school anti-bullying rules

BUFFALO, N.Y. ? Ryan Lysek rose to become vice president of his fifth-grade class at Lorraine Academy in Buffalo, New York. The sitting veep got bounced for saying things that went against the school's anti-bullying rules. So the 10-year-old is a little puzzled that candidates running to lead the co...

How It Works

Analyzer

Challenge

Go to Analyzer

All Grade Levels

Challenge Level



NASA's Valkyrie robots set the table for human life on Mars

LOWELL, Mass. ? Four sister robots built by NASA could be pioneers in the colonization of Mars, part of an advance construction team that sets up a habitat for more fragile human explorers. But first they're finding new homes on Earth and engineers to hone their skills. The space agency has kept on...

Grade Level: Original

Target Measure: 12.51

Target Measure: 12.40



L.A. subway work burps up a bonanza of prehistoric swag

LOS ANGELES? Seventy feet below Wilshire Boulevard, subway construction crews were hard at work. Fresh air roaring from giant ventilation pipes partly covered up the sickly sweet smell of petroleum. Amid the clatter of jackhammers and the whine of a mini-excavator, paleontologist Kim Scott scouted ...

Grade Level: 9--10th

Target Measure: 12.73



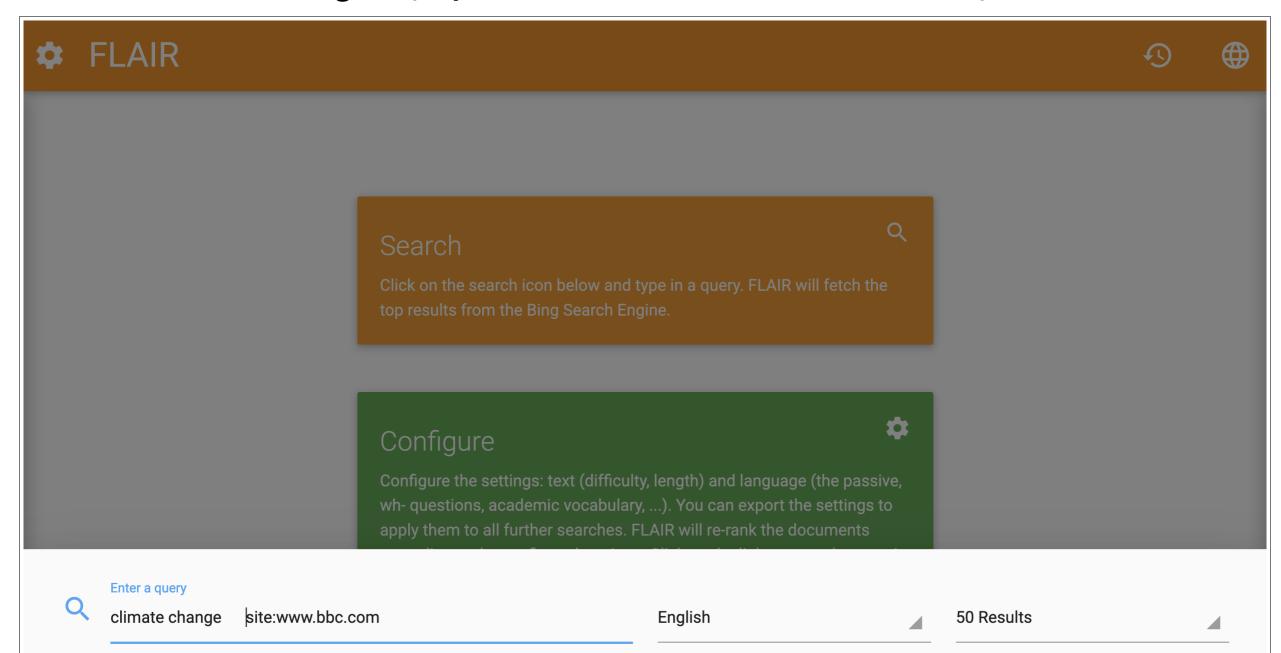
Don?t count on dance class to give your kid a workout, study warns

Only 8 percent of kids enrolled in after-school dance programs got enough exercise during class to meet federal guidelines for moderate to vigorous physical activity, a new study finds. Even worse, only 6 percent of adolescents got the recommended 30 minutes of moderate to vigorous exercise during t...

From research prototype to a tool usable in practice

- Exclusive focus on specific linguistic complexity aspect is too academic for real-life use.
- How about supporting selection of material that is
- rich in developmentally proximal language
 - → **Teachers** specify general proficiency level and pedagogically-focused language means.
- motivating in terms of content
 - → **Learners** search for texts that fit their individual interests.
- ⇒ FLAIR Search Engine (http://flair.schule, Chinkina & Meurers 2016)

FLAIR Search Engine (http://flair.schule, Chinkina & Meurers 2016)



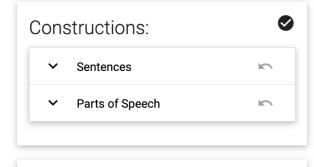


50 Results (0 Filtered)



SHARE SEARCH SETUP <





Language Use:

Academic Vocabulary

'climate change site:www.bbc.com'

1 What is climate change? A really simple guide - BBC News https://www.bbc.com/news/science-environment-24021772

Climate change is a shift in those average conditions. The rapid climate change we are now seeing is caused by humans using oil, gas and coal for their homes, factories and transport. When...

A brief history of climate change - BBC News
https://www.bbc.com/news/science-environment-15874560

Climate change hardly registers on the agenda, which centres on issues such as chemical pollution, atomic bomb testing and whaling. The United Nations Environment Programme (Unep) is formed as a ...

Climate change: IPCC report warns of 'irreversible' impacts ... - ...
https://www.bbc.com/news/science-environment-60525591

The Intergovernmental Panel on Climate Change says that humans and nature are being pushed beyond their abilities to adapt. Over 40% of the world's population are "highly vulnerable" to...

Climate change: 12 years to save the planet? Make that 18 mo...

https://www.bbc.com/news/science-environment-48964736

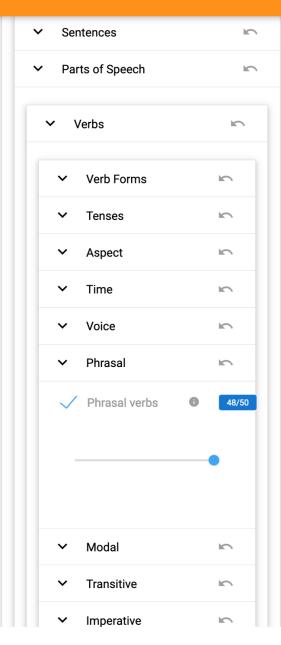
Prince Charles has stressed how important the next 12 months are in tackling climate change One of the understated headlines in last year's IPCC report was that global emissions of carbon dioxide...

Climate change: Young people very worried - survey - BBC News

Q







'climate change site:www.bbc.com'

I he law that could make climate change illegal - BBC Future

https://www.bbc.com/future/article/20200706-the-law-that-could-make-climate-change-illegal

Climate change is a global problem – if it is not tackled everywhere, it will affect everyone. The modern world is also incredibly intertwined: products – and resultan emissions – made in one...

Climate change: US formally withdraws from Paris agreement https://www.bbc.com/news/science-environment-54797743

After a three-year delay, the US has become the first nation in the world to formally withdraw from the Paris climate agreement. President Trump announced the move in June 2017, but UN regulations ...

Climate change: Can an enormous seaweed farm help curb it?

https://www.bbc.com/news/science-environment-63200589

He's confident his Seafields floating farm will draw sufficient CO2 from the air to moderate the effects of climate change, while also earning its backers carbon credits. At 55,000 sq km (21,200 ...

Plant-based diet can fight climate change - UN - BBC News https://www.bbc.com/news/science-environment-49238749

Climate change poses a threat to the security of our food supply. Rising temperatures, increased rain and more extreme weather events will all have an impact on crops and livestock.

The climate films shaping society - BBC Future
https://www.bbc.com/future/article/20221021-how-climate-change-in-film-shape-society

The law that could make climate chan... X

B1-B2 149 Sentence(s) 2802 Word(s)

2020. People are trustrated, but several governments have come and gone since the goal was set. "Don't blame us," the current government says. "We didn't take the decisions that led us here.".

The short-term cycles of government can be a real problem for climate change. Even if climate goals are laid down in law, there can often be few concrete measures to stop a succession of governments from taking decisions that collectively end up with them being missed.

But a new and ambitious climate law recently passed in

Denmark tries to find a way around this problem, and some of the other common pitfalls of climate laws. It makes Denmark one of a small number of countries beginning to provide new blueprints of how government can genuinely tackle climate change. Its law could turn out to be one of the closest things yet to a law that would make climate change – or at least the lack of effort to stop it – genuinely illegal. You might also like:

?

Construction Count Weight

Verl

Verbs > Phrasal > Phrasal...

14

(5)

ALL CONSTRUCTIONS

Export document



Include text formatting



Facilitating SLA research using such ICALL approaches

- Linguistic complexity (Vedder, Kuiken & Michel 2019) & alignment research (Kim & Michel 2023)
- Do language learners show alignment with increasingly complex ("i+1") language?
- Chen & Meurers (2019): analyzed alignment in continuation writing (data by Wang & Wang 2015)
 - ⇒ 403 of 575 linguistic complexity measures show alignment
- Chen, Meurers & Rebuschat (2022): ICALL continuation writing study
 - 112 Chinese learners of English
 - worked on sequence of ten increasingly complex texts, with four different challenge groups (MLTU)
 - ⇒ Learners in low and middle challenge group showed more alignment than high challenge group.
 - ⇒ There is alignment, but no clear evidence for learning.
 - Some future directions:
 - What constitutes + 1 in the different dimensions of linguistic complexity?
 - When does alignment lead to lasting learning effects?

Key Factors: **Noticing** and **Focus-on-Form**

- Input alone only results in Basic Varieties (Klein & Perdue 1997).
- Some attention to forms crucial: Noticing Hypothesis (Schmidt 1995)
- Foreign language learners need a Focus-on-Form (Long 1991; Lightbown 1998).
- increase salience (Sharwood Smith 1993)
- ICALL tools can support automated Input Enhancement (IE):
 - Visual: Highlighting of target forms on **user-selected** web pages (Meurers et al. 2010; Holz et al. 2018)
 - Functional: Generation of questions on sentences containing target forms (Chinkina & Meurers 2017)

Visual Input Enhancement example



Africa Americas Asia-Pacific Europe **Middle East South Asia**

England Northern Ireland Scotland Wales

Education

UK Politics

Magazine

Business Health

Science & **Environment Technology Entertainment**

E-mail this to a friend



Printable version

Cows also 'have regional accents'

Cows have regional accents like humans, language specialists have suggested.

They decided to examine the issue after dairy farmers noticed their cows had slightly different moos, depending on which herd they came from.

John Wells, Professor of Phonetics at the University of London, said regional twangs had been seen before in birds.



Cows moo with a regional twang



Listen Cow moo recordings

The farmers in Somerset who noticed the phenomenon said it may have been the result of the close bond between them and their animals.

Farmer Lloyd Green, from Glastonbury, said: "I spend a lot of time with my ones and they definitely moo with a Somerset drawl.

SEE ALSO

- 'Accent' confirms unique species 15 Aug 06 | Highlands and Islands
- Brain bug changes woman's accent 10 Jul 06 | Staffordshire
- What makes you local? 18 Feb 05 | Magazine

RELATED INTERNET LINKS

University of London phonetics department

The BBC is not responsible for the content of external internet sites

TOP UK STORIES

- Major manhunt for Afghan soldier
- Unemployment dips to 2.47 million
- PM condemns sympathy for Moat



News feeds

MOST POPULAR STORIES NOW

MOST SHARED

MOST READ

1 BBC News

Source: http://news.bbc.co.uk/2/hi/5277090.stm

Visual Input Enhancement example: making prepositions salient



Africa Americas Asia-Pacific Europe **Middle East** South Asia

England Northern Ireland Scotland Wales **UK Politics** Education Magazine

Health Science & **Environment Technology Entertainment**

Business

E-mail this to a friend



Printable version

Cows also 'have regional accents'

Cows have regional accents like humans, language specialists have suggested.

They decided to examine the issue after dairy farmers noticed their cows had slightly different moos, depending on which herd they came from.

John Wells, Professor of Phonetics at the University of London, said regional twangs had been seen before in birds.



Cows moo with a regional twang



Listen Cow moo recordings

The farmers in Somerset who noticed the phenomenon said it may have been the result of the close bond between them and their animals.

Farmer Lloyd Green, from Glastonbury, said: "I spend a lot of time with my ones and they definitely moo with a Somerset drawl.

SEE ALSO

- 'Accent' confirms unique species 15 Aug 06 | Highlands and Islands
- Brain bug changes woman's accent 10 Jul 06 | Staffordshire
- What makes you local? 18 Feb 05 | Magazine

RELATED INTERNET LINKS

University of London phonetics department

The BBC is not responsible for the content of external internet sites

TOP UK STORIES

- Warnings **before** Potters Bar crash
- Gaza ship Britons 'to head home'
- Walcott left out of England squad



MOST POPULAR STORIES NOW

MOST SHARED

MOST READ

1 Doubt cast on Israel raid

Source: http://news.bbc.co.uk/2/hi/5277090.stm

Facilitating SLA research using such ICALL approaches

- Ziegler et al. (2017): exploratory study using ICALL to support noticing of determiners
- Ruiz, Rebuschat & Meurers (2021): Study targeting the acquisition of English phrasal verbs
 - phrasal verbs are frequent and important for interpretation, but notoriously difficult (\rightarrow avoidance)
 - 127 participants (advanced L2 English, L1 German) read 40 texts in a two week span
 - randomly assigned to two conditions: meaning-focused and form-focused
 - form-focused group saw phrasal verb visually highlighted and particle as multiple choice item

EBERHARD KARLS





Business

Markets

World

UK

Tech

Money

Commentary

Breakingviews

Sport

Life

Pictures

First read the article, then click the button at the end of the article to move on to the questionnaire.

RETIREMENT NEWS | Mon Aug 15, 2016 | 9:34am EDT

Are you saving too much for your kids' college?

By Chris Taylor I NEW YORK

NEW YORK Parents these days are expected to pull off a financial Mission: Impossible. Cover the monthly bills, pay around \$ X @ debts, help elderly parents, save for retirement and for kids' college costs, all with incomes that may have been flat for years.

Facing such a demanding feat, here's some advice you may not often hear: When it comes to your kids' college costs, maybe you are doing too much.

To wit, 42 percent of parents are actually losing sleep over college costs, up from 28 percent just two years ago, according to new data from the Parents, Kids & Money survey by Baltimore-based money managers T. Rowe Price.

Of parents surveyed, 57 percent are willing to tak
√
are willing to borrow \$100,000 or more.

More parents (58 percent) report having college-s

"Parents are more stressed than ever about colleg

at least \$25,000 of college debt on behalf of their kids, and 19 percent

nts for their kids, rather than retirement savings for themselves (54 percent).

feel guilty about not being able to help more, and many are willing to take

off

down through

Facilitating SLA research using such ICALL approaches

- Ziegler et al. (2017): exploratory study using ICALL to support noticing of determiners
- Ruiz et al. (2021): Study targeting the acquisition of English phrasal verbs
 - phrasal verbs are frequent and important for interpretation, but notoriously difficult (\rightarrow avoidance)
 - 127 participants (advanced L2 English, L1 German) read 40 texts in a two week span
 - randomly assigned to two conditions: meaning-focused and form-focused
 - form-focused group saw phrasal verb visually highlighted and particle as multiple choice item

ICALL setup integrated assessment of working memory (Ruiz et al. 2019):

- OSpan (Unsworth et al. 2005), but using Klingon (Hicks et al. 2016) to reduce cheating
- interaction effect between working memory and form-focused instruction
 - ⇒ more benefit from attention to form for learners with higher working memory capacity
 - WM assists simultaneous processing of form, meaning and use of language forms (Williams 2012)
 - confirms relevance and feasibility of investigating interaction effects (DeKeyser 2012)

EBERHARD KARLS

TÜBINGEN



Key Factors: **Practice**, **Feedback**, **ZPD**

- Skill Acquisition Theory (DeKeyser 2005)
 - Teaching provides declarative knowledge, then **practice** is needed for automated skills.
 - Goal: fluent ability that does not bind attention
- Support learners in their individual Zone of Proximal Development (Vygotsky 1986)
 - → scaffolded feedback while working on an activity: micro-adaptivity
 - → adaptive selection of learning materials/activities: macro-adaptivity

Very challenging for teachers in practice: lack of time, diagnostics, adaptive materials

- ⇒ ICALL tools can adaptively support individualized practice (Ruiz, Rebuschat & Meurers 2024)
 - Petersen (2010): ICALL dialog system Sascha to practice English question formation
 - Intelligent Tutoring System FeedBook (Rudzewitz et al. 2017; Meurers et al. 2019)

March 19, 2024

Facilitating SLA research using such ICALL approaches

- Petersen (2010): English question acquisition
 - research context: developmental sequences proposed by Pienemann & Johnston (1987)
 - research question: effectiveness of recasts in spoken and written mode
- The ICALL system Sascha developed by Ken Petersen supports dialog interaction for information gap, story sequencing, and spot the difference activities.

Session 1 : Spot the Differences : Picture Set 2-A Student 1111 08:32 Is the bird on his shoulder? ASK SASHA

Facilitating SLA research using such ICALL approaches

- Petersen (2010): supporting acquisition of questions in English
 - research context: developmental sequences proposed by Pienemann & Johnston (1987)
 - research question: effectiveness of recasts in spoken and written mode
- The ICALL system Sascha developed by Ken Petersen supports dialog interaction for information gap, story sequencing, and spot the difference activities.
- Study design:
 - 56 ESL at public high school, Washington DC, mostly Spanish L1 (13 novice, 43 intermediate)
 - three groups: human-computer (ICALL), human-human, control
 - intervention: three times 30 minutes interactions with Sascha system or human interlocutor
- ⇒ Result: ICALL users showed development along expected sequences, parallel to human-human oral interaction, both better than control.

EBERHARD KARLS

Hilfreich?

○Ja ○Nein



From printed workbook to Al-based FeedBook

Grammar check: Problems Everyone has got problems. What could these people do differently? AUSGA	ANGS-
0. Gillian is sad. Her mother never has any time for her.	
If Mrs Collins had more time for Gillian, Gillian wouldn't be so sad.	
1. Mrs Collins feels bad. She should listen more to Gillian. The ske listen more to Gillian. She feels to	A Herri
2. Gwynn is very disappointed. Gillian doesn't like Wildings School as much as I If Gillian like Wildings School as much of	Welcome to Wales
	CYP2 Grammar check: Problems Everyone has got problems. What could these people do differently? Use conditional clauses in your answers.
	o. Gillian is sad. Her mother never has any time for her. If Mrs. Collins had more time for Gillian, Gillian wouldn't be so sad.
	Mrs Collins feels bad. She should listen more to Gillian. If she listens more to Gillian, she feels better. With conditional clauses (type 2) we use the
	2. Gwynn is very disappointed. Gillian doesn't like Wilding with conditional clauses (type 2), we use the simple past in the if-clause, not the simple present.

3. George and Rajiv feel bad because they don't have a pres

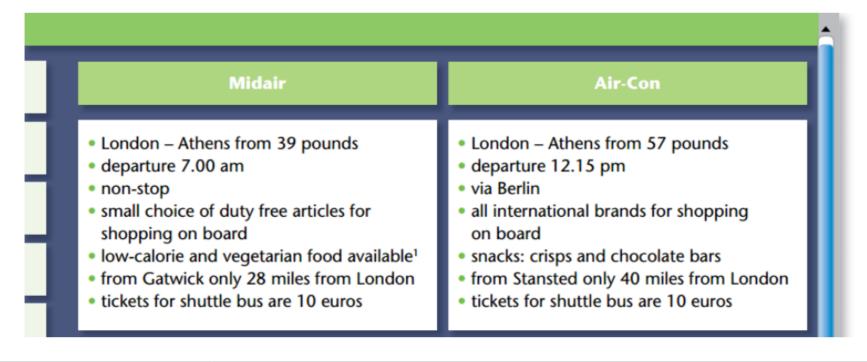
FeedBook: Feedback on language forms

Off to Greece again

Mr Lambraki is checking flights to Greece. Read the information he has found on the two airlines and use the adjectives below to compare them.

LiF8R: Comparison of adjectives

expensive (ticket) · early (departure) · attractive (shopping on board) · good (choice of food offered on board) · healthy (food and drinks) · suitable (airport) · cheap (tickets for shuttle bus) · friendly (service on board) · easy (online booking)

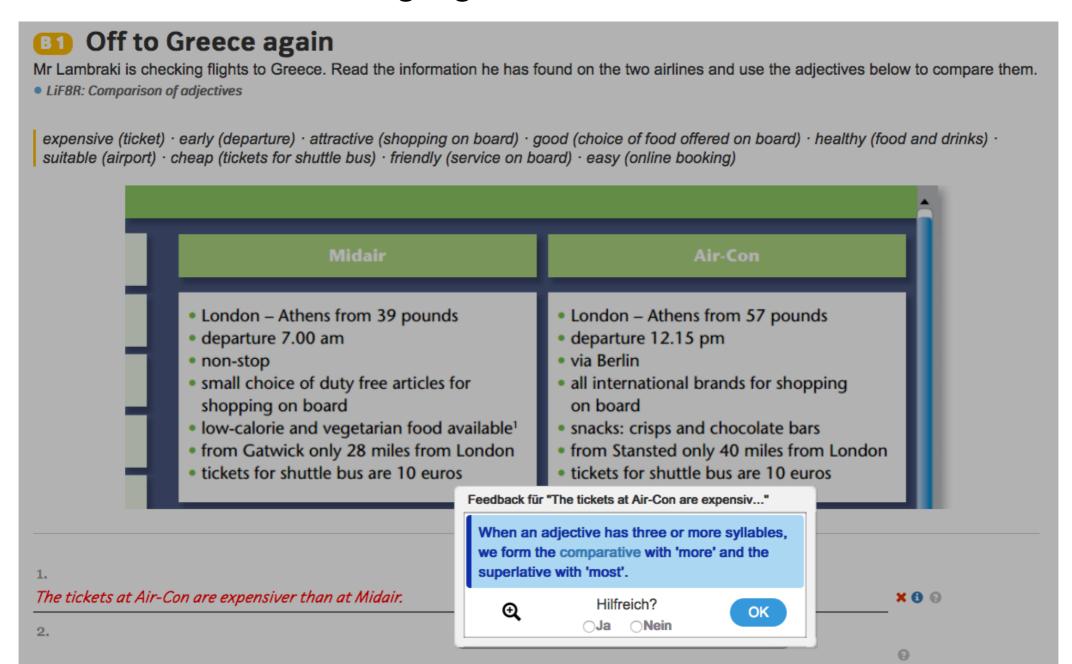


1.

The tickets at Air-Con are expensiver than at Midair.

9

FeedBook: Feedback on language forms



Gillian's diary

Read Gillian's diary entry and complete these sentences.

Friday 23rd September

We'll drive to the north coast tomorrow and have a look at two boarding schools with Gwynn. His sister went to Wildings and he says it would be great for me. If they had a football team, it wouldn't be so bad, but it's all so girly-girly with horses and ballet dancing and everything, yuk! I don't want to go to the school in Llandysul either. It still feels like Gwynn and Mum just want me to go to boarding school because of the new baby. Miss my pals like crazy, miss London and my old school. If my friends were here, boarding school would actually be fun. Being the new girl at school without any friends will be horrible. And it's all Gwynn's fault! The village where we live now is the worst. 20 minutes to the nearest supermarket. If there was a shopping

centre, I could at least go shopping. But there's nothing, no shops, no cinema, no nothing ... only sheep!

My room is really nice and big though and we have a fab garden which is great for playing football. If I made some friends in the village, we could have a great time there. I met Gruffudd, the boy from next door, this afternoon. He seems nice. He started talking to me in Welsh and I couldn't understand ANYTHING. He then spoke English and told me he plays rugby. Well, it's not football but I might have to learn to like it. They all love rugby here.

25

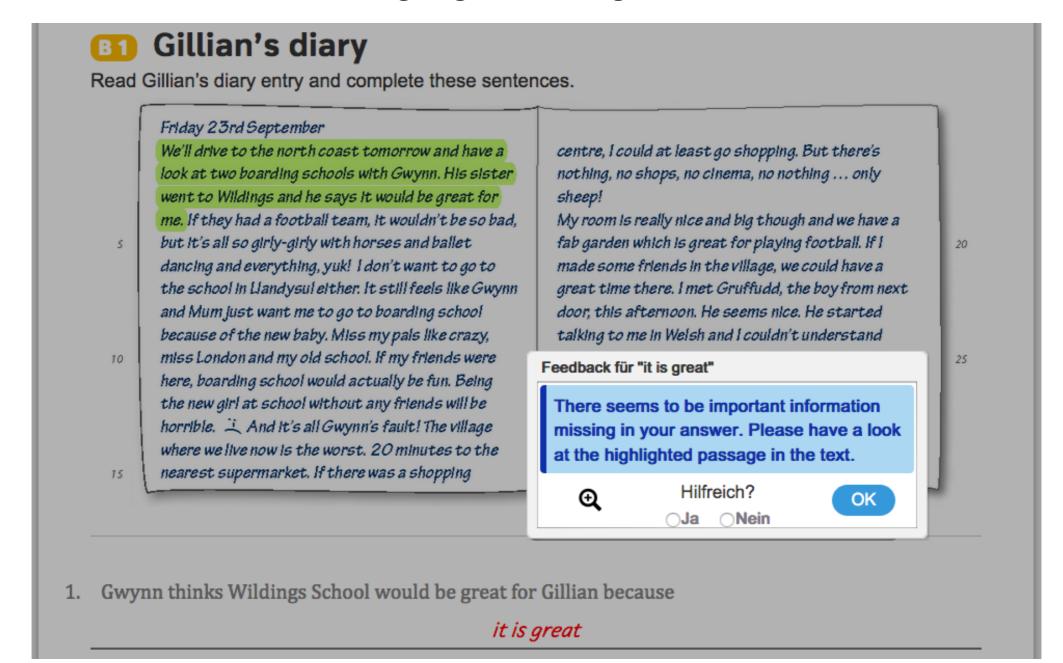
Aargh, If I had some credit on my phone, I could call Caroline. Hope I get some pocket money tomorrow.

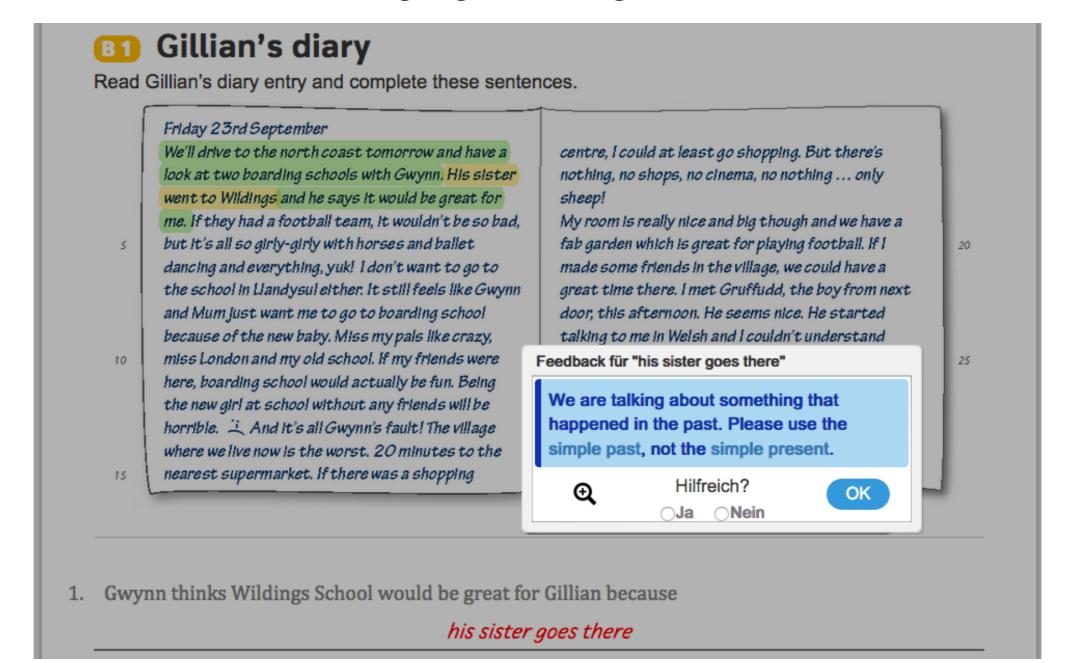
1. Gwynn thinks Wildings School would be great for Gillian because

it is great

10

15





Gillian's diary

Read Gillian's diary entry and complete these sentences.

Friday 23rd September

We'll drive to the north coast tomorrow and have a look at two boarding schools with Gwynn. His sister went to Wildings and he says it would be great for me. If they had a football team, it wouldn't be so bad, but it's all so girly-girly with horses and ballet dancing and everything, yuk! I don't want to go to the school in Llandysul either. It still feels like Gwynn and Mum just want me to go to boarding school because of the new baby. Miss my pals like crazy, miss London and my old school. If my friends were here, boarding school would actually be fun. Being the new girl at school without any friends will be horrible. And it's all Gwynn's fault! The village where we live now is the worst. 20 minutes to the nearest supermarket. If there was a shopping

centre, I could at least go shopping. But there's nothing, no shops, no cinema, no nothing ... only sheep!

My room is really nice and big though and we have a fab garden which is great for playing football. If I made some friends in the village, we could have a great time there. I met Gruffudd, the boy from next door, this afternoon. He seems nice. He started talking to me in Welsh and I couldn't understand ANYTHING. He then spoke English and told me he plays rugby. Well, it's not football but I might have to learn to like it. They all love rugby here.

20

25

Aargh, if I had some credit on my phone, I could call Caroline. Hope I get some pocket money tomorrow.

1. Gwynn thinks Wildings School would be great for Gillian because

his sister went there

0

35

Facilitating SLA research using such ICALL approaches

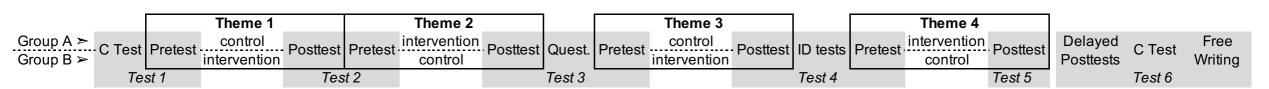
- The FeedBook covers all of the pedagogical language targets of the 7th grade English curriculum in Germany → can support full-year studies
 - 188 metalinguistic feedback types formulated by teachers (on tenses, comparatives, relative & reflexive pronouns, gerunds, passives, conditionals, (in)direct speech).
- Research question: Does immediate scaffolded feedback on form during homework improve learning of the targeted language means? (Meurers et al. 2019)
 - Subjects: Ten regular English classes in 7th grade academic track schools (Gymnasium), including a CLIL class, a private school, an all-boys class, and teachers of different generations
 - Setup: Regular teaching, but FeedBook as workbook.

EBERHARD KARLS

FeedBook as a platform for randomized field trials

- How can a system be fully embedded so that it supports regular teaching and support a randomized, controlled study?
- everyone uses the FeedBook as workbook (avoids temporary multimedia effect) and receives positive feedback and feedback on meaning, orthography, and default feedback
- within-class randomization: who sees specific feedback on which grammar topics
 - Overall study design:

EBERHARD KARLS

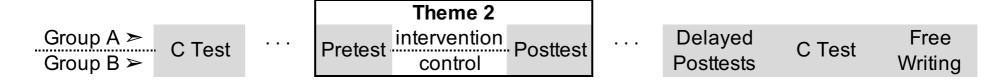


Each pre-/posttest targets the grammar topics of one Theme.

Zooming in on specific results

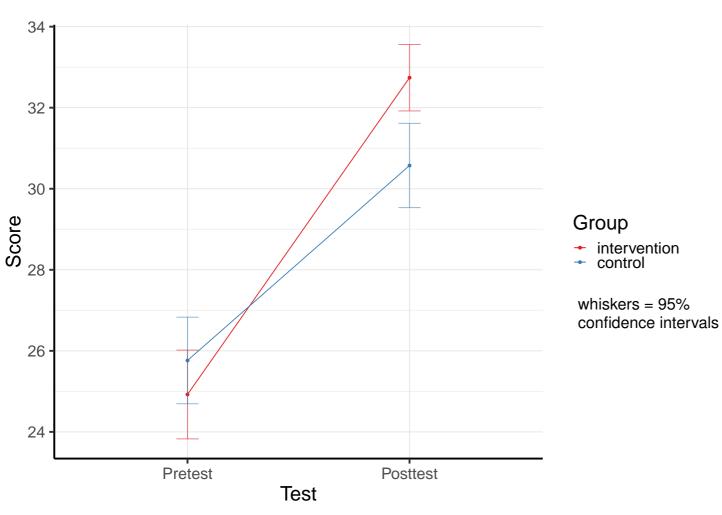
EBERHARD KARLS

Analysis of Theme 2 data (Meurers et al. 2019):



- Phenomena covered in Theme 2: comparatives, conditional clauses, relative clauses
- 205 students completed pre- and posttest for Theme 2.

Results on effectiveness of scaffolded feedback: pre-/posttest scores



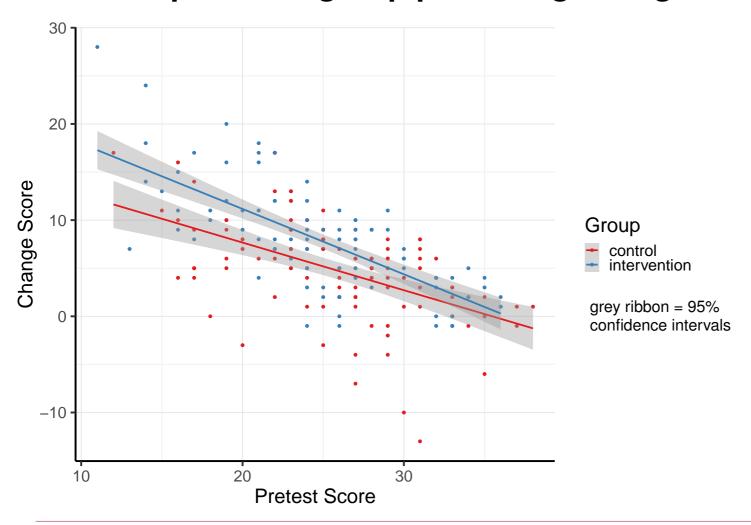
- Intervention group improved significantly more than control (p < 0.0001, Welsh two sample t-test):
 - 7.82 for intervention group
 - 4.81 for control group
- → intervention group learned 62% more than control group
 - Effect size: Cohen's d = 0.56

EBERHARD KARLS

TÜBINGEN



Results: pretest & group predicting change score



- "Worse" students improved more
- Possibly a ceiling effect of measuring instrument (max. score: 40)

EBERHARD KARLS

TÜBINGEN

Std. Error

Results: mixed effects logistic regression

- Is impact of feedback detectable independent of learner, item, class?
- Predict log odds of the binary outcome of each gap in the tests.
 - Fixed effects: test (pre/post), group (intervention/control) & their interaction
 - Random intercepts: test items, learners, teachers (=classes)

Results:	Estimat	
	(Intercept) = pretest control	0 73***

(Intercept) = pretest control	0.73***	(0.15)
TestPosttest	0.68***	(0.05)
GroupIntervention	-0.07	(0.09)
TestPosttest:GroupIntervention	0.43***	(0.08)
Var: Learner (Intercept)	0.30	
Var: Item (Intercept)	0.52	
Var: Teacher (Intercept)	0.05	

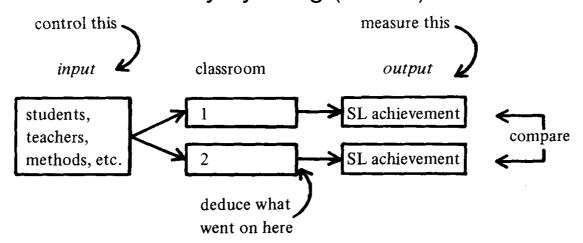
^{***}p < 0.001, **p < 0.01, *p < 0.05

28/43

EBERHARD KARLS

FeedBook as a platform for randomized field trials II

- Full integration into the regular school context = rely only on the state curriculum
 - Single requirement on teacher: assign at least two exercises for each language focus in curriculum
- Black box problem pointed out already by Long (1980b):



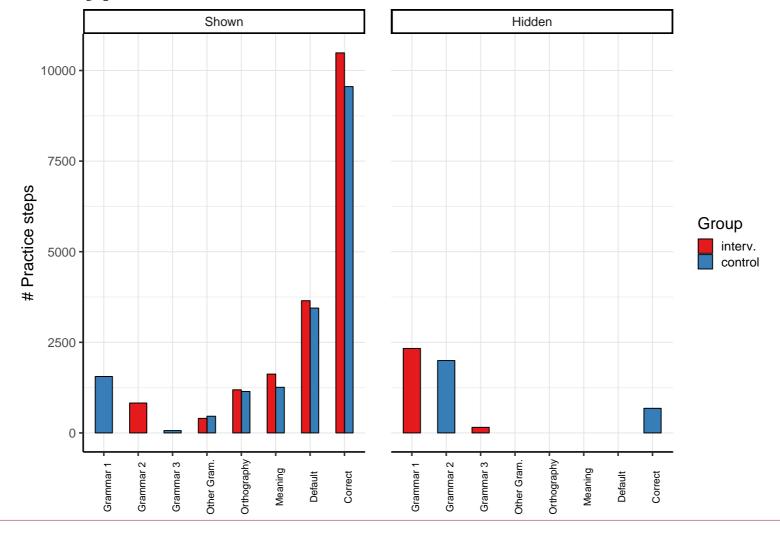
Problem: Loss of control resulting from scaling up to real-life school context.

Learning analytics using system log data shows us what the students actually did!

EBERHARD KARLS



Who saw which type of feedback?



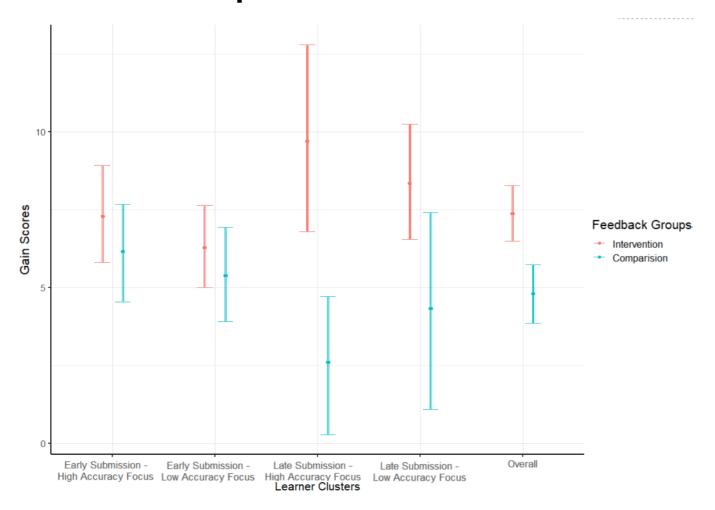
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Learning analytics: For whom is micro-adaptivity effective?

- How are learning gains and interaction patterns linked? (Hui, Rudzewitz & Meurers 2023)
- Clustering identifies four groups of students, based on:
 - +/- accuracy focus of student How many of the items were filled out correctly, not left empty, answered correctly at first try?
 - +/- submission time of student relative to peers
- ⇒ Scaffolded feedback has the biggest effect for students who
 - systematically attempt to solve the exercises correctly
 - submit later than their peers



Who benefits most from the specific feedback?



EBERHARD KARLS

TÜBINGEN

Learning analytics complements but does not replace learning product data

- Does the effect of practice of specific language forms transfer to open, functional language production? (De Kuthy & Meurers 2022)
- Writing topic chosen so that practiced language means are relevant:
 - Write a text about your holidays. Please include the following aspects:
 - Compare two of your holiday trips (weather, duration, ...),
 - describe your next holiday trip, and
 - outline what you would do if you could spend 1000€ during your next holiday.
- Automatic identification of practiced language forms in the learner writing
- ⇒ The score on the form-specific delayed posttest is a significant predictor of the number of uses of those forms in the free writing!
 - But the effect is driven only by *comparatives* due to challenge to determine when something should count as a relative or conditional clause. \rightarrow acquisition criterion?

Learning products as Big Data for SLA research

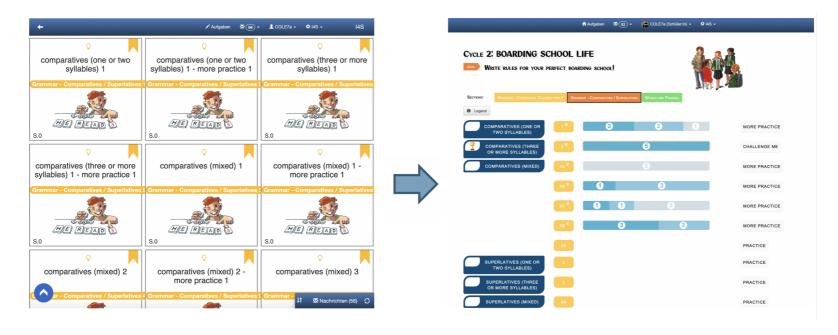
- System use in real life contexts also facilitates large-scale collection of learner product data.
- EFCamDat corpus based on Englishtown system: 1.2 million writings by 175,000 learners
 - individual variation and role of L1 in morpheme acquisition order Murakami (2013, 2016)
 - task-effects on linguistic complexity (Alexopoulou et al. 2017; Michel et al. 2019)
- Opportunities arising from "Big Data" in SLA (Alexopoulou, Meurers & Murakami 2022)

EBERHARD KARLS



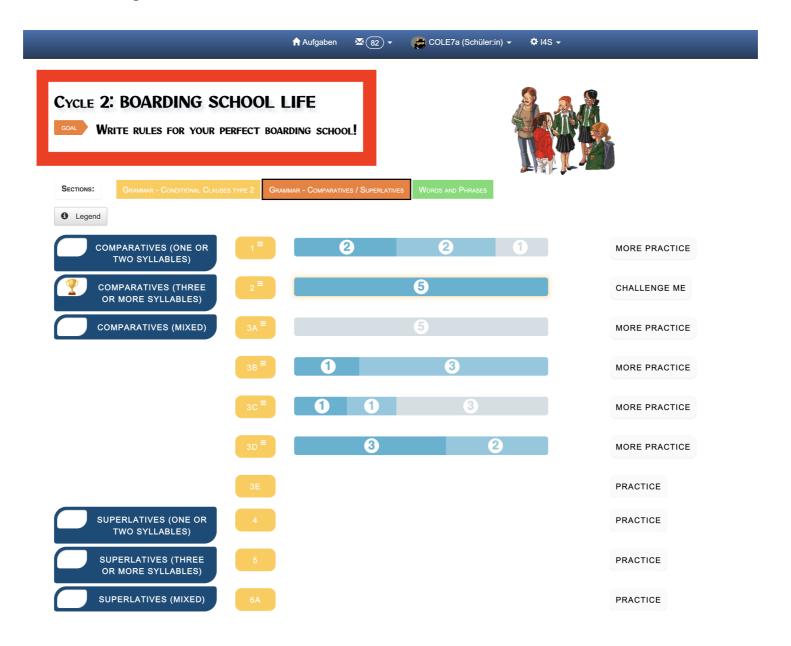
Integration of individual practice & teacher-orchestrated task-based class?

- Organize exercises so they empower students to successfully participate in functional tasks.
 - improves acceptance of practice as pre-task activities in a task-based curriculum
 - fosters intrinsic motivation and self-awareness of ability for students

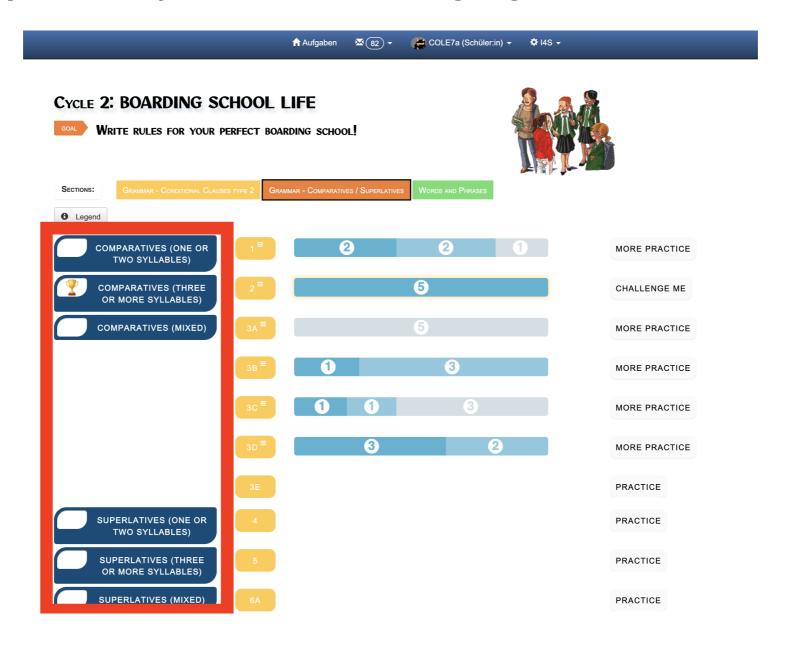


EBERHARD KARLS

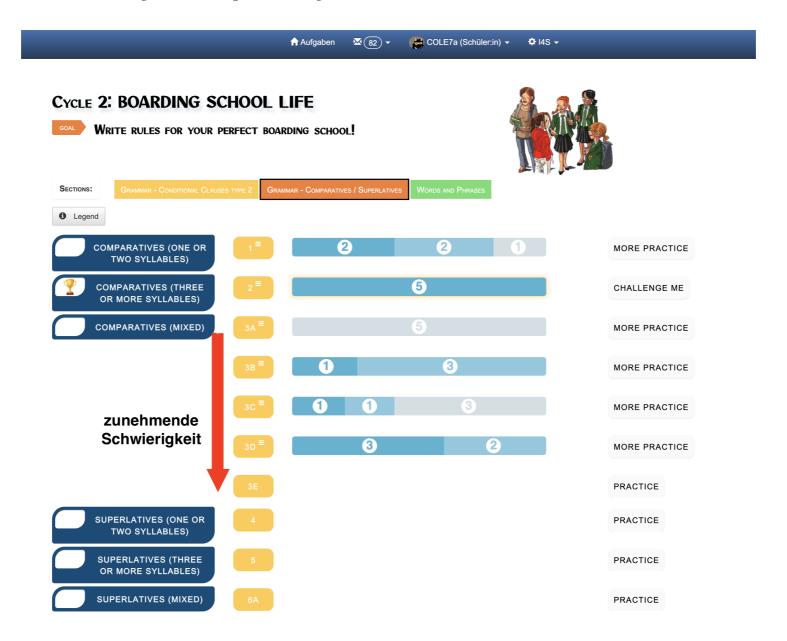
1. Make final task explicit



2. Organize practice by task-essential language means



3. Incremental activity complexity



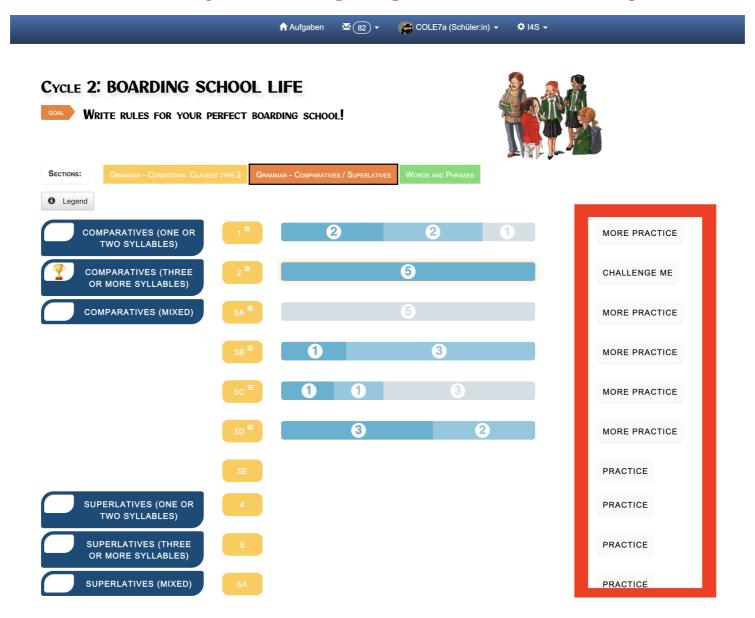
4. Feedback on success relative to support received



5. Feedback on acquisition criterion



7. Selection of next activity via language means to be practiced



First results on Interact4School RCT using task-oriented FeedBook

- 847 students from 36 classes (English 7th grade high school)
- broad range of analyses in preparation, incl. classroom videography by partners at Univ. of Lüneburg: Torben Schmidt, Diana Pili-Moss, Carolyn Blume
- Focus here: 618 students in FeedBook condition (7 schools, 24 classes)
- ⇒ Analysis of learning success (pretest/posttest) shows significantly higher learning gains, when comparing new dashboard version to original FeedBook (Parrisius et al. submitted)
 - student questionnaire data also confirm perceived helpfulness of student dashboard
- → Interestingly, motivational feedback through avatar was detrimental after initial phase.

EBERHARD KARLS

43/43

Next steps in randomized controlled field trials

- First study: scaffolded feedback (Meurers et al. 2019)
- Second study: task-orientation with criterial and motivational feedback (Parrisius et al. 2022)
- Third study: individually adaptive activity sequencing integrating ambulatory assessment of working memory (http://digbindiff.de)
- Next study: Support an interface to teacher-orchestrated classroom, requiring teacher dashboard and teacher training in SLA and pedagogy (http://ai2teach.de)
 - requires further scaling up to have enough power for testing variables at teacher-level
- Related research: Facilitating spoken language interactive dialogues: AISLA Project (Xiaobin Chen)

Summary

- Digital tools using AI methods can help realize mechanisms established by SLA research in educational practice
 - improve real-life language teaching and learning
 - support further refinement of SLA concepts and hypotheses in authentic, ecologically-valid contexts
- This opens up possibilities for conducting SLA research in real-life settings:
 - web-based intervention studies supporting scalable analysis of interaction effects
 - randomized controlled field studies in authentic school setting, supporting analysis of product and process data (learning analytics)
- Pursuing these opportunities requires interdisciplinary collaboration on
 - SLA concepts and hypotheses to be refined and tested
 - Al methods to analyze and generate language, model learners, activities and curricular learning goals
 - joint analysis of large-scale process and product data

References

- Alexopoulou, T., D. Meurers & A. Murakami (2022). Big Data in SLA: advances in methodology and analysis. In N. Ziegler & M. Gonzalez-Lloret (eds.), Routledge Handbook of Second Language Acquisition and Technology, Routledge. URL http://purl.org/dm/papers/Alexopoulou.Murakami.Meurers-22.pdf.
- Alexopoulou, T., M. Michel, A. Murakami & D. Meurers (2017). Task Effects on Linguistic Complexity and Accuracy: A Large-Scale Learner Corpus Analysis Employing Natural Language Processing Techniques. Language Learning 67, 181–209. URL https://doi.org/10.1111/lang.12232.
- Aljaafreh, A. L. & J. P. Lantolf (1994). Negative feedback as regulation and second language learning in the zone of proximal development. The Modern Language Journal 78(4), 465–483.
- Chen, X. (2019). Automatic analysis of linguistic complexity and its application in language learning research. Ph.D. thesis, Universität Tübingen.
- Chen, X. & D. Meurers (2017). Challenging Learners in Their Individual Zone of Proximal Development Using Pedagogic Developmental Benchmarks of Syntactic Complexity. In *Proceedings of the Joint 6th Workshop on NLP for Computer Assisted* Language Learning and 2nd Workshop on NLP for Research on Language Acquisition at NoDaLiDa 2017. Gothenburg, Sweden: ACL, Linköping Electronic Conference Proceedings 134, pp. 8–17. URL http://aclweb.org/anthology/W17-0302.pdf.
- Chen, X. & D. Meurers (2019). Linking text readability and learner proficiency using linguistic complexity feature vector distance. Computer-Assisted Language Learning 32(4), 418–447. URL https://doi.org/10.1080/09588221.2018.1527358. https://doi.org/10.1080/09588221.2018.1527358.

References II

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

- Chen, X., D. Meurers & P. Rebuschat (2022). ICALL offering individually adaptive input: Effects of complex input on L2 development. Language Learning & Technology 26(1). URL https://hdl.handle.net/10125/73496.
- Chinkina, M. & D. Meurers (2016). Linguistically-Aware Information Retrieval: Providing Input Enrichment for Second Language Learners. In Proceedings of the 11th Workshop on Innovative Use of NLP for Building Educational Applications (BEA). San Diego, CA: ACL, pp. 188–198. URL http://aclweb.org/anthology/W16-0521.pdf.
- Chinkina, M. & D. Meurers (2017). Question Generation for Language Learning: From ensuring texts are read to supporting learning. In Proceedings of the 12th Workshop on Innovative Use of NLP for Building Educational Applications (BEA). Copenhagen, Denmark, pp. 334–344. URL http://aclweb.org/anthology/W17-5038.pdf.
- Corder, S. P. (1967). The Significance of Learner's Errors. *International Review of Applied Linguistics in Language Teaching* 5(4), 161–170.
- De Kuthy, K. & D. Meurers (2022). Extending experimental research on the effectiveness of an Intelligent Tutoring System. A corpus study of authentic ESL student essays. In *Learner Corpus Research Conference*. Book of Abstracts. Padua, Italy.
- DeKeyser, R. M. (2005). What Makes Learning Second-Language Grammar Difficult? A Review of Issues. Language Learning 55(S1), 1–25.
- DeKeyser, R. M. (2012). Interactions Between Individual Differences, Treatments, and Structures in SLA. Language Learning 62, 189-200.
- Hicks, K. L., J. L. Foster & R. W. Engle (2016). Measuring working memory capacity on the web with the Online Working Memory Lab (the OWL). Journal of Applied Research in Memory and Cognition 5, 478–489.

References III

- Holz, H., Z. Weiss, O. Brehm & D. Meurers (2018). COAST Customizable Online Syllable Enhancement in Texts. A flexible framework for automatically enhancing reading materials. In *Proceedings of the Thirteenth Workshop on Innovative Use of NLP for* Building Educational Applications. New Orleans, Louisiana: Association for Computational Linguistics, pp. 89–100. URL https://aclanthology.org/W18-0509.
- Hui, B., B. Rudzewitz & D. Meurers (2023). Learning Processes in Interactive CALL systems: Linking Automatic Feedback, System Logs, and Learning Outcomes. Language Learning and Technology 27(1), 1–23. URL https://hdl.handle.net/10125/73527.
- Kim, Y. & M. Michel (2023). Linguistic alignment in second language acquisition: A methodological review. *System* p. 103007.
- Klein, W. & C. Perdue (1997). The Basic Variety (or: Couldn't natural languages be much simpler?). Second language research 13(4), 301-347.
- Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. New York: Longman.
- Larsen-Freeman, D. (2000). Second Language Acquisition and Applied Linguistics. *Annual Review of Applied Linguistics* 20, 165-181.
- Lightbown, P. M. (1998). The importance of timing in focus on form. In C. Doughty & J. Williams (eds.), Focus on Form in Classroom Second Language Acquisition, Cambridge: Cambridge Univ. Press, pp. 177–196.
- Long, M. H. (1980a). Input, interaction, and second language acquisition. Ph.D. thesis, University of California, Los Angeles.
- Long, M. H. (1980b). Inside the "black box": Methodological issues in classroom research on language learning. Language Learning 30(1), 1–42.

References IV

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

- Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. De Bot, C. Kramsch & R. Ginsberg (eds.), Foreign language research in cross-cultural perspective, Amsterdam: John Benjamins, pp. 39–52.
- Meurers, D., K. De Kuthy, V. Möller, F. Nuxoll, B. Rudzewitz & R. Ziai (2018). Digitale Differenzierung benötigt Informationen zu Sprache, Aufgabe und Lerner. Zur Generierung von individuellem Feedback in einem interaktiven Arbeitsheft. FLuL – Fremdsprachen Lehren und Lernen 47(2), 64–82. URL http://purl.org/dm/papers/Meurers.DeKuthy.ea-18.pdf.
- Meurers, D., K. De Kuthy, F. Nuxoll, B. Rudzewitz & R. Ziai (2019). Scaling up intervention studies to investigate real-life foreign language learning in school. Annual Review of Applied Linguistics 39, 161–188.
- Meurers, D., R. Ziai, L. Amaral, A. Boyd, A. Dimitrov, V. Metcalf & N. Ott (2010). Enhancing Authentic Web Pages for Language Learners. In *Proceedings of the 5th Workshop on Innovative Use of NLP for Building Educational Applications (BEA)*. Los Angeles: ACL, pp. 10–18. URL http://aclweb.org/anthology/W10-1002.pdf.
- Michel, M., A. Murakami, T. Alexopoulou & D. Meurers (2019). Effects of task type on morphosyntactic complexity across proficiency: Evidence from a large learner corpus of A1 to C2 writings. *Instructed Second Language Acquisition* 3, 124–152. URL https://doi.org/10.1558/isla.38248.
- Murakami, A. (2013). Individual Variation and the Role of L1 in the L2 Development of English Grammatical Morphemes: Insights From Learner Corpora. Ph.D. thesis, University of Cambridge.
- Murakami, A. (2016). Modeling Systematicity and Individuality in Nonlinear Second Language Development: The Case of English Grammatical Morphemes. Language Learning 6(4), 834–871.

References V

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

- Parrisius, C., I. Pieronczyk et al. (2022). Using an intelligent tutoring system within a task-based learning approach in English as a foreign language classes to foster motivation and learning outcome (Interact4School): Pre-registration of the study design. PsychArchives .
- Parrisius, C., K. Wendebourg et al. (eingereicht). Effective features of feedback in an intelligent tutoring system. *Language Learning* and Technology .
- Petersen, K. (2010). Implicit Corrective Feedback in Computer-Guided Interaction: Does Mode Matter? Ph.D. thesis, Georgetown University. URL http://purl.org/net/Petersen-10.pdf.
- Pica, T. & C. Doughty (1985). The role of group work in classroom second language acquisition. Studies in second language acquisition 7(2), 233-248.
- Pienemann, M. (1984). Psychological constraints on the teachability of languages. Studies in Second Language Acquisition 6, 186–214. URL http://www.neiu.edu/~circill/bofman/ling460/psychological.pdf.
- Pienemann, M. & M. Johnston (1987). Factors influencing the development of language proficiency. In D. Nunan (ed.), *Applying* second language acquisition research, Adelaide: National Curriculum Resource Centre, Adult Migrant Education Program.
- Robinson, P. (2003). The Cognition Hypothesis, Task Design, and Adult Task-Based Language Learning. Second Language Studies 21, 45–107.
- Rudzewitz, B., R. Ziai, K. De Kuthy & D. Meurers (2017). Developing a web-based workbook for English supporting the interaction of students and teachers. In Proceedings of the Joint 6th Workshop on NLP for Computer Assisted Language Learning and 2nd Workshop on NLP for Research on Language Acquisition. pp. 36–46. URL http://aclweb.org/anthology/W17-0305.pdf.

49/43

References VI

- Ruiz, S., X. Chen, P. Rebuschat & D. Meurers (2019). Measuring individual differences in cognitive abilities in the lab and on the web. PloS one 14(12), e0226217.
- Ruiz, S., P. Rebuschat & D. Meurers (2021). The effects of working memory and declarative memory on instructed second language vocabulary learning: Insights from intelligent CALL. Language Teaching Research 25(4), 510–539.
- Ruiz, S., P. Rebuschat & D. Meurers (2024). Supporting Individualized practice through Intelligent CALL. In Y. Suzuki (ed.), *Practice* and automatization in second language research: Perspectives from skill acquisition theory and cognitive psychology, Routledge, pp. 119–143. URL https://purl.org/dm/papers/Ruiz.Rebuschat.Meurers-24.pdf.
- Ruiz Hernández, S. E. (2018). Individual differences and instructed Second Language Acquisition: Insights from Intelligent Computer Assisted Language Learning. Doctoral dissertation, Eberhard-Karls-Universität Tübingen.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics* 11, 206–226. URL http://applij.oxfordjournals.org/cgi/reprint/11/2/129.pdf.
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (ed.), Attention and awareness in foreign language learning, Honolulu, HI: University of Hawaii, pp. 1–63.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics* 10(3), 209–231.
- Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. Studies in Second Language Acquisition 15, 165-179. URL https://doi.org/10.1017/s0272263100011943.
- Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (ed.), *Handbook on research in second language* teaching and learning, Mahwah, NJ: Lawrence Erlbaum Associates, pp. 471–484.

References VII

- Unsworth, N., R. P. Heitz, J. C. Schrock & R. W. Engle (2005). An automated version of the operation span task. *Behavior Research* Methods 37(3), 498-505.
- Vedder, I., F. Kuiken & M. Michel (2019). Linguistic complexity and instruction in second language acquisition: introduction. *Instructed* Second Language Acquisition 3(2), 119–123. URL https://journal.equinoxpub.com/ISLA/article/view/13912.
- Vygotsky, L. S. (1986). Thought and Language. Cambridge, MA: MIT Press.
- Wang, C. & M. Wang (2015). Effect of alignment on L2 written production. *Applied Linguistics* 36(5), 503–526.
- Weiss, Z., S. Dittrich & D. Meurers (2018). A Linguistically-Informed Search Engine to Identify Reading Material for Functional Illiteracy Classes. In Proceedings of the 7th Workshop on Natural Language Processing for Computer-Assisted Language Learning (NLP4CALL). URL https://aclweb.org/anthology/W18-7109.pdf.
- Williams, J. N. (2012). Working memory and SLA. In P. Robinson, A. Mackey, S. M. Gass & R. Schmidt (eds.), *The Routledge* handbook of second language acquisition, New York: Routledge, pp. 427–441.
- Ziegler, N., D. Meurers, P. Rebuschat, S. Ruiz, J. L. Moreno Vega, M. Chinkina, W. Li & S. Grey (2017). Interdisciplinary research at the intersection of CALL, NLP, and SLA: Methodological implications from an input enhancement project. Language Learning 67, 209-231.