Intelligent Computer-Assisted Language Learning

Authentic Text ICALL (ATICALL)
Exercise Generation & Language Aware Search Engine

Detmar Meurers
(Universität Tübingen)

based on joint research:

(Meurers et al. 2010; Ott & Meurers 2010; Vajjala & Meurers 2012)

Tromsø, 2. October 2013

Introduction

- The use of NLP in language learning has primarily centered on analyzing learner language in Tutoring Systems and Learner Corpora.
- How about using NLP of authentic native language in support of language learning?
  - WERTi/VIEW: Automatic visual input enhancement of web pages and generation of activities (Meurers et al. 2010)
  - LAWSE: Language-Aware Search Engine, retrieving authentic texts at the appropriate level for learners (Ott & Meurers 2010)
    - Determining readability of native texts based on SLA measures of development (Vajjala & Meurers 2012)

Grounding in SLA research (Schmidt 1995)

- Noticing
  - "conscious registration of an event"
- Understanding
  - "recognition of a general principle, rule or pattern"
- There is no learning without noticing, but awareness requires input.

⇒ Learners have to be exposed to linguistic features to acquire them and must notice them.

- Strategies highlighting the salience of language forms and categories are referred to as **input enhancement** (Sharwood Smith 1993).

⇒ Let's use NLP to provide automatic input enhancement for language learners!

WERTi: Working with English Real Text

- Provide learners of English (ESL) with input enhancement for any web pages they are interested in.

→ good for learner motivation:
  - learners can choose material based on their interests
  - includes news, up-to-date information, hip stuff
  - pages remain fully contextualized (video, audio, links)

→ wide range of potential learning contexts:
  - can supplement **regular classroom instruction**
  - can support voluntary, self-motivated pursuit of knowledge, i.e., **lifelong learning**.
  - can foster **implicit learning**, e.g., for adult immigrants:
    - already functionally living in second language environment, but stagnating in acquisition
    - without access/motivation to engage in explicit learning, but browsing the web for information and entertainment

What language properties should we enhance?

- A wide range of linguistic features can be relevant for **awareness**, incl. morphological, syntactic, semantic, and pragmatic information (Schmidt 1995).
- We focus on enhancing language patterns which are well-established difficulties for ESL learners:
  - determiner and preposition usage
  - use of gerunds vs. *to*-infinitives
  - *wh*-question formation
  - phrasal verbs

  NLP identifying other patterns can easily be integrated as part of a flexible NLP architecture.

How should the targeted forms be enhanced?

- WERTi currently offers three types of input enhancement:
  a) color **highlighting** of the pattern or selected parts thereof
  b) pages supporting **clicking**, with automatic color feedback
     - automatic feedback compares automatic annotation of clicked on form with targeted form
  c) pages supporting practice (e.g., **fill-in-the-blank**), with automatic color feedback
     - automatic feedback compares form entered by learner with form in original text
- This follows standard pedagogical practice ("PPP"):
  a) receptive presentation
  b) presentation supporting limited interaction
  c) controlled practice
  d) (free production)

Prepositions: Presentation (Color)

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

Prepositions: Practice (FIB)

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

Prepositions: Presentation + Interaction (Click)

Car-free cities: an idea with legs

Car-free neighbourhoods are no unrealistic utopia – they exist all over Europe

(55) (110)
Tweet this (121)
Comments (68)

'Not anti-car, just pro-choice' ... a cyclist in Vauban, Germany. Photograph: Sipa Press/Rex Features

A quarter of households in Britain – more in the larger cities, and a majority in some inner cities – live without a car. Imagine how quality of life would improve for cyclists and everyone else if traffic were removed from areas where people could practically choose to live without cars. Does this sound unrealistic, utopian? Did you know many European cities are already doing it?

Posted by Steve Melia Thursday 29 October 2009 08.00 GMT
guardian.co.uk

A larger | smaller

Life and style
Cycling

Environment
Ethical and green living · Travel and transport

Series
Bike blog

More from Green living blog on

Life and style
Cycling

Environment
Ethical and green living · Travel and transport

Source: http://www.guardian.co.uk/environment/green-living-blog/2009/oct/29/car-free-cities-neighbourhoods

Prepositions: Presentation + Interaction (Click)

Car-free cities: an idea with legs

Car-free neighbourhoods are no unrealistic utopia – they exist all over Europe

(55) (110)
Tweet this (121)
Comments (68)

'Not anti-car, just pro-choice' ... a cyclist in Vauban, Germany. Photograph: Sipa Press/Rex Features

A quarter of households in Britain – more in the larger cities, and a majority in some inner cities – live without a car. Imagine how quality of life would improve for cyclists and everyone else if traffic were removed from areas where people could practically choose to live without cars. Does this sound unrealistic, utopian? Did you know many European cities are already doing it?

Posted by Steve Melia Thursday 29 October 2009 08.00 GMT
guardian.co.uk

A

Life and style

Environment

Series

More from Green living blog on

Life and style

Environment

Series

More blogposts

Source: http://www.guardian.co.uk/environment/green-living-blog/2009/oct/29/car-free-cities-neighbourhoods

Phrasal verbs: Presentation (Color)

Laugh Lines

Funny Stuff From All Over

May 6, 2010, 11:14 AM

Letterman: 'They Don't Like Immigrants'

Monologue | *Wednesday night on "The Late Show With David Letterman" on CBS*: You folks been following the big British Petroleum oil spill in the Gulf of Mexico? I'm telling you, British Petroleum has put more birds in oil than Colonel Sanders.

I was thinking about this. Here's what I came up with. Now, in Arizona, you know about the new immigration law, where if you don't look like you belong there, they can run you out of the state? And they've got patrol cars driving around, pulling up to people, saying: "You don't look like you belong here. Get out!" So the deal is, in Arizona, they don't like immigrants. And I was thinking, well, that's odd, because right across the river there in California, they elected one governor.

E-mail This Print Share Twitter Sign in to Recommend

Phrasal verbs: Presentation + Interaction (Click)

Laugh Lines

Funny Stuff From All Over

May 6, 2010, 11:14 AM

Letterman: 'They Don't Like Immigrants'

Monologue | *Wednesday night on "The Late Show With David Letterman" on CBS*: You folks been following the big British Petroleum oil spill in the Gulf of Mexico? I'm telling you, British Petroleum has put more birds in oil than Colonel Sanders.

I was thinking about this. Here's what I came up with. Now, in Arizona, you know about the new immigration law, where if you don't look like you belong there, they can run you out of the state? And they've got patrol cars driving around, pulling up to people, saying: "You don't look like you belong here. Get out!" So the deal is, in Arizona, they don't like immigrants. And I was thinking, well, that's odd, because right across the river there in California, they elected one governor.

E-mail This Print Share Twitter Sign in to Recommend

Phrasal verbs: Practice (Fill-in-the-blank)

Laugh Lines

Funny Stuff From All Over

May 6, 2010, 11:14 AM

Letterman: 'They Don't Like Immigrants'

Monologue | *Wednesday night on "The Late Show With David Letterman" on CBS*: You folks been following the big British Petroleum oil spill in the Gulf of Mexico? I'm telling you, British Petroleum has **put** more birds [ ] [?] oil than Colonel Sanders.

I was thinking about this. Here's what I **came up** with. Now, in Arizona, you know about the new immigration law, where if you don't look like you belong there, they can **run** you [ ] [?] the state? And they've got patrol cars driving around, **pulling up** to people, saying: "You don't look like you belong here. **Get** **on** [?] !" So the deal is, in Arizona, they don't like immigrants. And I was thinking, well, that's odd, because right across the river there in California, they elected one governor.

E-mail This | Print | Share | Twitter | Sign in to Recommend

Gerunds vs. infinitives: Presentation (Color)

"The government says it is expanding access to university, but they are actually blocking people's aspirations and betraying a generation."

The government was forced to cap student numbers **after** **discovering** a £200m black hole in the university financing budget at the end of last year. Labour was accused **of** **abandoning** its pledge **to** **expand** higher education, addingpressure to a growing debate about how **to** **fund** the growing number of young people who **want** **to** **do** a degree. The government is due **to** **announce** a review of student finance.

The massive increase in applicants has put a strain on the university system this year, with one university forced **to** **convert** single bedrooms in halls into doubles, and others putting students up in hotels.

Source: http://www.guardian.co.uk/education/2009/oct/14/30000-miss-university-place

Gerunds vs. infinitives: Practice (FIB)

"The government says it is expanding access to university, but they are actually blocking people's aspirations and betraying a generation."

The government was forced to cap student numbers after [ ] [?] *(discover)* a £200m black hole in the university financing budget at the end of last year. Labour was accused of [ ] [?] *(abandon)* its pledge to expand higher education, addingpressure to a growing debate about how to fund the growing number of young people who want [ ] [?] *(do)* a degree. The government is due to announce a review of student finance.

The massive increase in applicants has put a strain on the university system this year, with one university forced to convert single bedrooms in halls into doubles, and others putting students up in hotels.

Source: http://www.guardian.co.uk/education/2009/oct/14/30000-miss-university-place

Wh-questions: Presentation (Color)

If someone takes drugs, they can become addictive depending on the drug. Overdoses typically happen with cocaine, opioids, benzos, especially mixing benzos and opioids (Xanex, Valium, or Klonopin).

Why do people use illegal drugs? [change]

Most illegal drugs cause people to become intoxicated[*needs proving*]. The slang term for this experience is "getting stoned" or "getting high." When a drug user is intoxicated, they may feel strange, happy, dizzy, or weird. Some drugs such as marijuana and hashish often make users feel sleepy and relaxed. Some drug users have feelings that they are floating or dreaming. Drugs such as LSD make people feel intensely; they make one see and feel things like never before, and think things about the world they would normally not. Some say it increases knowledge and creates wisdom. Other drugs such as Crystal Meth make users feel excited and happy and full of energy.

Source: http://simple.wikipedia.org/wiki/Illegal_drugs

Wh-questions: Presentation + Interaction (Click)

If someone takes drugs, they can become addictive depending on the drug. Overdoses typically happen with cocaine, opioids, benzos, especially mixing benzos and opioids (Xanex, Valium, or Klonopin).

Why do people use illegal drugs? subject [change]

Most illegal drugs cause people to become intoxicated[needs proving]. The slang term for this experience is "getting stoned" or "getting high." When a drug user is intoxicated, they may feel strange, happy, dizzy, or weird. Some drugs such as marijuana and hashish often make users feel sleepy and relaxed. Some drug users have feelings that they are floating or dreaming. Drugs such as LSD make people feel intensely; they make one see and feel things like never before, and think things about the world they would normally not. Some say it increases knowledge and creates wisdom. Other drugs such as Crystal Meth make users feel excited and happy and full of energy.

Source: http://simple.wikipedia.org/wiki/Illegal_drugs

Wh-questions: Presentation + Interaction (Click)

If someone takes drugs, they can become addictive depending on the drug. Overdoses typically happen with cocaine, opioids, benzos, especially mixing benzos and opioids (Xanex, Valium, or Klonopin).

Why do people use illegal drugs? subject [change]

Most illegal drugs cause people to become intoxicated[needs proving]. The slang term for this experience is "getting stoned" or "getting high." When a drug user is intoxicated, they may feel strange, happy, dizzy, or weird. Some drugs such as marijuana and hashish often make users feel sleepy and relaxed. Some drug users have feelings that they are floating or dreaming. Drugs such as LSD make people feel intensely; they make one see and feel things like never before, and think things about the world they would normally not. Some say it increases knowledge and creates wisdom. Other drugs such as Crystal Meth make users feel excited and happy and full of energy.

Source: http://simple.wikipedia.org/wiki/Illegal_drugs

Wh-questions: Practice (FIB)

If someone takes drugs, they can become addictive depending on the drug. Overdoses typically happen with cocaine, opioids, benzos, especially mixing benzos and opioids (Xanex, Valium, or Klonopin).

illegal drugs? do people Why use [change]

Why people do use illegal drugs? ✓ ✗ ?

Most illegal drugs cause people to become intoxicated[needs proving]. The slang term for this experience is "getting stoned" or "getting high." When a drug user is intoxicated, they may feel strange, happy, dizzy, or weird. Some drugs such as marijuana and hashish often make users feel sleepy and relaxed. Some drug users have feelings that they are floating or dreaming. Drugs such as LSD make people feel intensely; they make one see and feel things like never before, and think things about the world they would normally not. Some say it increases knowledge and creates wisdom. Other drugs such as Crystal Meth make users feel excited and happy and full of energy.

Source: http://simple.wikipedia.org/wiki/Illegal_drugs

Wh-questions: Practice (FIB)

If someone takes drugs, they can become addictive depending on the drug. Overdoses typically happen with cocaine, opioids, benzos, especially mixing benzos and opioids (Xanex, Valium, or Klonopin).

Why do people use illegal drugs? [change]

Most illegal drugs cause people to become intoxicated[needs proving]. The slang term for this experience is "getting stoned" or "getting high." When a drug user is intoxicated, they may feel strange, happy, dizzy, or weird. Some drugs such as marijuana and hashish often make users feel sleepy and relaxed. Some drug users have feelings that they are floating or dreaming. Drugs such as LSD make people feel intensely; they make one see and feel things like never before, and think things about the world they would normally not. Some say it increases knowledge and creates wisdom. Other drugs such as Crystal Meth make users feel excited and happy and full of energy.

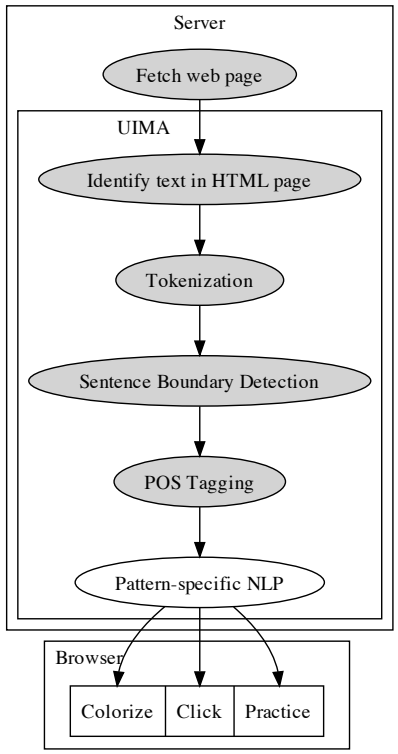
Source: http://simple.wikipedia.org/wiki/Illegal_drugs

Realizing WERTi

- Guiding ideas behind implementation:
  - Reuse existing NLP tools where possible
  - Support integration of a range of language patterns
- First WERTi prototype (Amaral/Meurers/Metcalf at CALICO 06, EUROCALL 06)
  - implemented in Python using NLTK (Bird & Loper 2004), TreeTagger (Schmid 1994)
  - integrated into Apache2 webserver using mod_python
  - input enhancement targets: determiners and prepositions in Reuters news text
  - still available at `http://purl.org/icall/werti-v1`
- How can we flexibly support integration of a wider range of language patterns using heterogeneous set of NLP?
  → integrate NLP into UIMA-based architecture on server

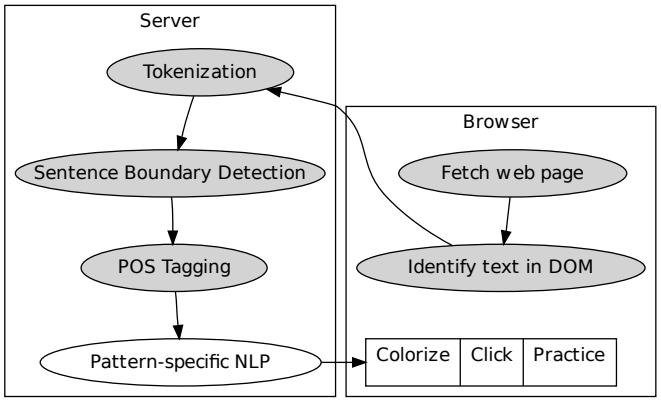
WERTi architecture

- reimplementation in Java (Dimitrov/Ziai/Ott)
- Tomcat servlet
- idea behind architecture
  - use same core processing
  - demand-driven pattern-specific NLP
- input enhancement targets:
  - determiners
  - prepositions
  - gerunds vs. *to*-infinitives
  - tense in conditionals
  - *wh*-questions



WERTi architecture: Browser plugin version

Firefox plugin (Adriane Boyd) moves fetching of web page and text identification to client to better support sites requiring login, cookies, or dynamically generated text.



- http://purl.org/icall/werti
- international VIEW version: http://purl.org/icall/view

Pattern-specific NLP

- UIMA-based architecture (Ferrucci & Lally 2004)
  - each NLP tool **annotates** the input
    - OpenNLP tools, LingPipe tagger, TreeTagger, Constraint Grammar CG 3
  - UIMA data repository is common to all components (Götz & Suhre 2004)
- We use available pre-trained models for
  - TreeTagger with PennTreebank tagset
  - LingPipe Tagger with Brown tagset
  - OpenNLP tools (Tokenizer, Sentence Detector, Tagger, Chunker)
- Specify input enhancement targets
  - in terms of standard annotation schemes
    - e.g., identify determiners via `AT|DT|DTI|DTS|DTX` using Brown tagset
  - using constraint-grammar rules (CG 3 compiler), e.g.:
    - 101 rules for gerunds vs. *to*-infinitives
    - 126 rules for *wh*-question patterns

Evaluating input enhancement techniques

Does input enhancement improve learning outcomes?

- Improving learning outcomes is the overall goal of WERTi and visual input enhancement in general.
- While some studies show an improvement in learning outcomes, the study of visual input enhancement sorely needs more experimental studies (Lee & Huang 2008).
- WERTi can systematically produce visual input enhancement for a range of language properties
  - → Supports real-life foreign language teaching studies under a wide range of parameters.
  - → Supports lab-based experiments to evaluate when input enhancement succeeds in making learners notice enhanced properties (eye tracking, ERP).

Evaluating input enhancement techniques

High precision NLP needed for automatic input enhancement

- Automatic visual input enhancement requires reliable identification of the relevant classes using NLP.
  - Note: Precision of identification of specific classes relevant, not overall quality of POS-tagging or parsing.
- Problem 1: Often no established gold standard available for the language classes to be enhanced.
- Problem 2: Realistic test set must be established by studying what pages learners choose for enhancement.

Evaluating input enhancement techniques

Evaluating determiner and preposition identification

- Evaluation of preposition and determiner identification using BNC Sampler Corpus
  - high quality CLAWS-7 annotation provides gold standard for preposition and determiner classes
  - relatively broad representation of English
- Performance of the LingPipe POS tagger in WERTi:

	precision	recall
prepositions	95.07%	90.52%
determiners	97.06%	94.07%

Contextualizing our work

- NLP has received most attention in ICALL in connection with **analyzing learner language** to
  - provide feedback to the learner
  - guide learner through material according to performance
  - Note: Uses NLP to process learner language
- WERTi **analyzes native language texts** to
  - identify target language categories and forms to make learners aware of them and their context of use.
  - Note: Uses NLP to process well-formed, native language

  = Authentic Text ICALL (ATICALL)

Related work:

- Data-Driven Learning
- Automatic Exercise Generation
- Reading Support Tools

Related Work
Data-Driven Learning

- One can view automatic input enhancement as an enrichment of Data-Driven Learning (DDL).
  - DDL is an “attempt to cut out the middleman [the teacher] as far as possible and to give the learner direct access to the data” (Boulton 2009, p. 82, citing Tim Johns)
- WERTi: learner stays in control, but NLP uses ‘teacher knowledge’ about relevant language properties to make those more prominent to the learner.

Related Work
Automatic Exercise Generation

- Antoniadis et al. (2004) describes plans of MIRTO project to support “gap-filling” and “lexical spotting” exercises in combination with a corpus database.
- VISL project (Bick 2005) offers games and visual presentations to foster knowledge of syntactic forms/rules.
  - KillerFiller produces slot-filler exercises from corpus texts; presented in isolation, in a testing setup.

Related Work
Reading Support Tools

- Glosser-RuG (Nerbonne et al. 1998): supports reading of French texts for Dutch learners
  - context-dependent dictionary, morphological analysis, and examples of word use in corpora
- COMPASS project (Breidt & Feldweg 1997): similar to Glosser-RUG, focusing on multi-word lexemes
- ALPHEIOS project (http://alpheios.net): supports lexicon lookup and provides aligned translations
- REAP project (http://reap.cs.cmu.edu) supports learners in searching for texts that are well-suited for providing vocabulary and reading practice (Heilman et al. 2008).

Outlook: Questions to be addressed

- Which **language pattern types** should be input enhanced?
  - adverb placement
  - tense and aspect
    - while effect is semantic, lexical cues can be identified by NLP (“*usually* go” vs. “are going *tomorrow*”)
  - passive vs. active
  - . . .
- Which **aspect of the patterns** should be input enhanced?
  - lexical classes, morphemes
  - contextual clues (optional or obligatory)
- What is the **best input enhancement**, i.e., highlighting or interaction possibilities
  - for a particular linguistic pattern,
  - given a specific web page with its existing visual design features?

Finding texts appropriate for language learners

- How can one find authentic texts as reading material or for activity generation (e.g., WERTi)?
- Such texts should
  - be in the language of interest
  - have the appropriate level of complexity for the learner
  - contain enough good instances of the language patterns and rules targeted by the activities.
- How about simply using the web and a standard web search engine (e.g., google)?
  - Pro: The Web is huge, and up-to-date information on virtually any topic is available.
  - Cons: Standard search engines are not aware of reading complexity and language patterns.

$\Rightarrow$ Create a dedicated search engine for language learning!

Readability and how to measure it

- Readability or text difficulty refers to the understandability or comprehensibility of a text (Klare 1963).
- The more reading proficient the reader, the less readable texts need to be in order to be understood by this reader.
- Traditional readability formulas try to measure the readability on a scale, e.g., the U.S. grade level scale.

Traditional Readability Formulas

- Over two hundred traditional readability formulas have been developed (cf. Dubey 2004).
- They are generally developed for special purposes, such as determining the complexity of military training manuals (Caylor et al. 1973).
- A frequently used traditional readability measure is the Flesch-Kincaid formula (Kincaid et al. 1975)

Example: Flesh-Kincaid

- Computes U.S. grade level needed to read a text.
- Derived empirically from set of hand-classified documents.

$$\text{Flesch-Kincaid} = -15.59 + 11.8 \cdot AWL_s + 0.39 \cdot ASL$$

Where

$AWL_s = \frac{\text{Number of Syllables}}{\text{Number of Words}}$ Average word length counted in syllables.

$ASL = \frac{\text{Number of Words}}{\text{Number of Sentences}}$ Average sentence length.

- Idea:
  - The longer the word, the harder it is.
    (and the less common it is, cf. Zipf 1936)
  - The longer the sentence, the harder it is to understand.

Traditional readability measures: Evaluation

- Pros:
  - Relatively simple to use.
  - 'Simple' NLP only: tokenizer, stemming, sentence splitter, sometimes syllable counter
- Cons:
  - Originally developed and validated using very small and often highly specific data sets (e.g., technical manuals).
  - Whether the automated analysis agrees with the original human analysis has generally not been validated.
  - Measures such as sentence length are domain-dependent.
  - Underlying assumptions (e.g., 'long sentences are difficult') are rather crude generalizations.

Idea: Use SLA features for Readability

- Explore measures from SLA research
  - In SLA, measures of the complexity of learner writing are well-established as measures of acquisition process.
  - ⇒ We show that the best **measures of SLA development** are highly predictive for **readability classification** as well.
- Extend the most widely used WeeklyReader corpus by combining it with BBC-Bitesize corpus.
  - ⇒ Larger, more diverse corpus *WeeBit* covering more age groups is available for future research

Corpora: WeeklyReader

- Most previous CL research based on publicly accessible data used the graded reading material from the WeeklyReader web site.
- We there collected the following **WeeklyReader** corpus:

Grade Level	*Age in Years*	*Number of Articles*	*Avg. Number of Sentences/Article*
Level 2	7–8	629	23.41
Level 3	8–9	801	23.28
Level 4	9–10	814	28.12
Senior	10-12	1325	31.21

Corpora: WeeklyReader

- **WeeBit** corpus combines the **WeeklyReader** data with data from the **BBC Bitesize** website:

Grade Level	*Age in Years*	*Number of Articles*	*Avg. Number of Sentences/Article*
Level 2	7–8	629	23.41
Level 3	8–9	801	23.28
Level 4	9–10	814	28.12
KS3	11–14	644	22.71
GCSE	14–16	3500	27.85

Features

- Lexical Richness Features from SLA research (SLALEX) (Lu 2012)
  - Lexical Density measures: Various transformations of Type-Token Ratio
  - Lexical Variation measures: Proportion of words of various POS categories
- Syntactic complexity features from SLA research (SLASYN) (Lu 2010)
  - Mean length of different production units (sentence, clause, T-unit)
  - Sentence complexity: e.g., number of clauses per sentence
  - Coordination: coordinate phrases per clause, t-unit
  - Subordination: e.g., dependent clauses per clause, t-unit
  - Ratio of specific syntactic structures per production unit e.g., complex nominals per clause

More Features

- Other syntactic features
  - Average number of NPs, VPs, PPs, SBARs per sentence
  - Average number of dependent clauses, co-ordinate phrases, complex T-units per sentence
  - Mean length of NP, VP and PP
  - Average parse tree height
- Traditional features
  - Traditional Features (avg. word length, sentence length)
  - Traditional Formulae (Flesch-Kincaid score, Coleman-Liau score)
  - List of words (Academic Word List)

Best Features

- The BEST10FEATURES set was determined automatically using WEKA's Information Gain attribute selection method:
  - *dependent clause to clause ratio*
  - *complex nominals per clause*
  - *modifier variation* (proportion of adjectives and adverbs)
  - *adverb variation* (proportion of adverbs)
  - *mean length of a sentence*
  - avg. num. characters per word
  - avg. num. syllables per word
  - proportion of words in Academic Word List
  - num. co-ordinate phrases per sentence
  - Coleman-Liau Score
- Half of the best features are from SLA research (italics).

Experimental Setup

- Experiments used 500 training & 125 testing documents per level.
- Classifier used: Multi-Layer Perceptron (MLP) from WEKA toolkit.
- For syntactic features: Berkeley Parser and Tregex pattern matcher
- Included replication of syntactic feat. of Petersen & Ostendorf (2009)

Results

	Number of Features	Performance Accuracy	RMSE
Previous Work on WeeklyReader			
Feng (2010)	122	74%	
Petersen & Ostendorf (2009)	25	63.2%	
P. & O. syntactic features only	4	50.91%	
Our Results on WeeklyReader			
Replication P. & O. syntactic feat.	4	50.68%	
Our Syntactic Features	25	64.3%	0.37
Our Lexical Features	19	84.1%	0.23
All our Features	46	**91.3%**	0.17
Our Results on WeeBit			
Traditional Features	3	70.3%	0.25
SLALEX	16	68.1%	0.29
SLASYN	14	71.2%	0.28
SLALEX + SLASYN	30	**82.3%**	0.23
All Syntactic Features	25	75.3%	0.27
All Lexical Features	19	86.7%	0.20
All Features	46	**93.3%**	0.15
BEST10FEATURES	10	89.7%	0.18

Summary

- We motivated and discussed an approach providing automatic input enhancement of authentic web pages.
  - NLP identifies relevant linguistic categories and forms.
  - The sentences turned into activities can remain fully contextualized as part of the pages selected by learner.
- Automatic feedback for the practice activities is feasible since the original text is known.
  - Next step: Where possible alternatives exist, determine equivalence classes automatically; e.g., for prepositions building on Elghafari, Meurers & Wunsch (2010).
- Web pages can be selected by learners based on interests.
  - Language Aware Search Engine (Ott & Meurers 2010) can taking into account
    - content of interest to learner
    - language properties to be input enhanced
    - readability measures
  - Measures of development from SLA research are good predictors for readability classification (Vajjala & Meurers 2012)

References

Amaral, L., V. Metcalf & D. Meurers (2006). Language Awareness through Re-use of NLP Technology. Presentation at Preconference Workshop on "NLP in CALL – Computational and Linguistic Challenges". CALICO 2006. May 17, 2006. University of Hawaii. URL http://purl.org/net/icall/handouts/calico06-amaral-metcalf-meurers.pdf.

Antoniadis, G., S. Echinard, O. Kraif, T. Lebarbé, M. Loiseau & C. Ponton (2004). NLP-based scripting for CALL activities. In L. Lemnitzer, D. Meurers & E. Hinrichs (eds.), *Proceedings of eLearning for Computational Linguistics and Computational Linguistics for eLearning, International Workshop in Association with COLING 2004.*. Geneva, Switzerland: COLING, pp. 18–25. URL http://aclweb.org/anthology/W04-1703.

Bick, E. (2005). Grammar for Fun: IT-based Grammar Learning with VISL. In P. Juel (ed.), *CALL for the Nordic Languages*, Copenhagen: Samfundslitteratur, Copenhagen Studies in Language, pp. 49–64. URL http://beta.visl.sdu.dk/pdf/CALL2004.pdf.

Bird, S. & E. Loper (2004). NLTK: The Natural Language Toolkit. In *Proceedings of the ACL demonstration session*. Barcelona, Spain: Association for Computational Linguistics, pp. 214–217. URL http://aclweb.org/anthology/P04-3031.

Boulton, A. (2009). Data-driven Learning: Reasonable Fears and Rational Reassurance. *Indian Journal of Applied Linguistics* 35(1), 81–106.

Breidt, E. & H. Feldweg (1997). Accessing Foreign Languages with COMPASS. *Machine Translation* 12(1–2), 153–174. URL http://www.springerlink.com/content/v833605061168351/fulltext.pdf. Special Issue on New Tools for Human Translators.

Caylor, J. S., T. G. Sticht, L. C. Fox & J. P. Ford. (1973). *Methodologies for determining reading requirements of military occupational specialties: Technical report No. 73-5*. Tech. rep., Human Resources Research Organization, Alexandria, VA.

Dubey, A. (2004). Statistical Parsing for German: Modeling Syntactic Properties and Annotation Differences. Ph.D. thesis, Universität des Saarlandes. URL http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=FB3C44E0E4723F2011C09A210DD4A2AA?doi=10.1.1.69.5394&rep=rep1&type=pdf.

Elghafari, A., D. Meurers & H. Wunsch (2010). Exploring the Data-Driven Prediction of Prepositions in English. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING)*. Beijing, China, pp. 267–275. URL http://aclweb.org/anthology/C10-2031.

Ellis, N. (1994). Implicit and Explicit Language Learning - An Overview. In N. Ellis (ed.), *Implicit and explicit learning of languages*, London: Academic Press, pp. 1–31.

Feng, L. (2010). Automatic Readability Assessment. Ph.D. thesis, City University of New York (CUNY). URL http://lijun.symptotic.com/files/thesis.pdf?attredirects=0.

Ferrucci, D. & A. Lally (2004). UIMA: An architectural approach to unstructured information processing in the corporate research environment. *Natural Language Engineering* 10(3–4), 327–348. URL http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=252253&fulltextType=RA&fileId=S1351324904003523.

Götz, T. & O. Suhre (2004). Design and implementation of the UIMA Common Analysis System. *IBM Systems Journal* 43(3), 476–489. URL http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.90.5824&rep=rep1&type=pdf.

Heilman, M., L. Zhao, J. Pino & M. Eskenazi (2008). Retrieval of Reading Materials for Vocabulary and Reading Practice. In *Proceedings of the Third Workshop on Innovative Use of NLP for Building Educational Applications (BEA-3) at ACL'08*. Columbus, Ohio, pp. 80–88.

Johns, T. (1994). From printout to handout: Grammar and vocabulary teaching in the context of data-driven learning. In T. Odlin (ed.), *Perspectives on Pedagogical Grammar*, Cambridge: Cambridge University Press, pp. 293–313.

Kincaid, J. P., R. P. J. Fishburne, R. L. Rogers & B. S. Chissom (1975). *Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel*. Research Branch Report 8-75, Naval Technical Training Command, Millington, TN.

Klare, G. R. (1963). *The Measurement of Readability*. Ames, Iowa: Iowa State University Press.

Lee, S.-K. & H.-T. Huang (2008). Visual Input Enhancement and Grammar Learning: A Meta-Analytic Review. *Studies in Second Language Acquisition* 30, 307–331.

Lightbown, P. M. & N. Spada (1999). *How languages are learned*. Oxford: Oxford University Press. URL http://books.google.com/books?id=wlYTbuCsR7wC&lpg=PR11&ots=_DdFbrMCOO&dq=How20languages20are20learned&lr&pg=PR11#v=onepage&q&f=false.

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.

Long, M. H. (1996). The role of linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (eds.), *Handbook of second language acquisition*, New York: Academic Press, pp. 413–468.

Lu, X. (2010). Automatic analysis of syntactic complexity in second language writing. *International Journal of Corpus Linguistics* 15(4), 474–496.

Lu, X. (2012). The Relationship of Lexical Richness to the Quality of ESL Learners' Oral Narratives. *The Modern Languages Journal* pp. 190–208.

Lyster, R. (1998). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classroom. *Language Learning* 48, 183–218.

Metcalf, V. & D. Meurers (2006). Generating Web-based English Preposition Exercises from Real-World Texts. URL http://purl.org/net/icall/handouts/eurocallSchulz06-metcalf-meurers.pdf. Presentation at EUROCALL 2006. Granada, Spain. September 4–7, 2006.

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-5) at NAACL-HLT 2010*. Los Angeles: Association for Computational Linguistics, pp. 10–18. URL http://aclweb.org/anthology/W10-1002.pdf.

Nerbonne, J., D. Dokter & P. Smit (1998). Morphological Processing and Computer-Assisted Language Learning. *Computer Assisted Language Learning* 11(5), 543–559. URL http://urd.let.rug.nl/nerbonne/papers/call_fr.pdf.

Norris, J. M. & L. Ortega (2000). Effectiveness of L2 Instruction: A Research Synthesis and Quantitative Meta-Analysis. *Language Learning* 50(3), 417–528.

Authentic Text ICALL (ATICALL)
Detmar Meurers Universität Tübingen
Introduction
WERTi
What should we enhance?
How should it be enhanced?
Example activities
Realizing WERTi
Evaluation
Evaluating learning outcomes
Evaluating the NLP
Context
ICALL: ILTs and ATICALL
Data-driven learning
Exercise Generation
Reading Support Tools
Outlook
Language Aware Search Engine
Measuring Text Difficulty
SLA features for Readability
Corpora used
Features
Results
Summary

Ott, N. (2009). Information Retrieval for Language Learning: An Exploration of Text Difficulty Measures. ISCL master's thesis, Universität Tübingen, Seminar für Sprachwissenschaft, Tübingen, Germany. URL http://drni.de/zap/ma-thesis.

Ott, N. & D. Meurers (2010). Information Retrieval for Education: Making Search Engines Language Aware. *Themes in Science and Technology Education. Special issue on computer-aided language analysis, teaching and learning: Approaches, perspectives and applications* 3(1–2), 9–30. URL http://purl.org/dm/papers/ott-meurers-10.html.

Petersen, S. E. & M. Ostendorf (2009). A machine learning approach to reading level assessment. *Computer Speech and Language* 23, 86–106.

Schmid, H. (1994). Probabilistic Part-of-Speech Tagging Using Decision Trees. In *Proceedings of the International Conference on New Methods in Language Processing*. Manchester, UK, pp. 44–49. URL http://www.ims.uni-stuttgart.de/ftp/pub/corpora/tree-tagger1.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.

Schulz, R. A. (2002). Hilft es die Regel zu wissen um sie anzuwenden? Das Verhältnis von metalinguistischem Bewusstsein und grammatischer Kompetenz in DaF. *Die Unterrichtspraxis—Teaching German* 35(1), 15–24. URL http://www.jstor.org/stable/pdfplus/3531951.pdf.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition* 15, 165–179.

Vajjala, S. & D. Meurers (2012). On Improving the Accuracy of Readability Classification using Insights from Second Language Acquisition. In J. Tetreault, J. Burstein & C. Leacock (eds.), *In Proceedings of the 7th Workshop on Innovative Use of NLP for Building Educational Applications*. Montréal, Canada: Association for Computational Linguistics, pp. 163—-173. URL http://aclweb.org/anthology/W12-2019.pdf.

Zipf, G. K. (1936). *The Psycho-Biology of Language*. London: Routledge.

Authentic Text ICALL (ATICALL)
Detmar Meurers Universität Tübingen
Introduction
WERTi
What should we enhance?
How should it be enhanced?
Example activities
Realizing WERTi
Evaluation
Evaluating learning outcomes
Evaluating the NLP
Context
ICALL: ILTs and ATICALL
Data-driven learning
Exercise Generation
Reading Support Tools
Outlook
Language Aware Search Engine
Measuring Text Difficulty
SLA features for Readability
Corpora used
Features
Results
Summary