Language Learning and NLP: Connecting Needs and Opportunities

Detmar Meurers

Department of Linguistics and LEAD Graduate School
Eberhard-Karls Universität Tübingen

Kortrijk/Louvain-la-Neuve
21./22. April 2015

Introduction

- NLP so far plays no role in
  - real-life Foreign Language Teaching and Learning (FLTL)
  - or Second Language Acquisition (SLA) research.
- Are there actual needs for automatic language analysis in real-life FLTL practice or SLA research?
  - students in schools increasingly heterogeneous, requiring more individually adapted materials and interaction
  - life-long learning increasingly important
  - distance education requires more support for individual assessment and feedback, especially
    - when scaling up (e.g., MOOCs)
    - to support more open, functional tasks
  - large scale learner corpora are becoming available for SLA

⇒ NLP can play an important role addressing real needs

Introduction

NLP can address needs by supporting the analysis of

I. **learner language** to
   - analyze learner capabilities
   - provide interactive feedback
   - support interaction

II. **native language** to
   - search, modify, or enhance language input for learners
   - create exercise materials

that are adapted to individual learner needs

We here focus on analyzing written language; additional issues are involved in spoken language interaction and pronunciation training.

Why analyze learner language?

Intelligent Tutoring Systems (ITS)

- online analysis of learner language aimed at supporting language acquisition
- provide immediate, individualized scaffolding feedback:
  - meta-linguistic feedback in form-focused activities
  - incidental Focus-on-Form in meaning-based activities
  - feedback on meaning (essential for functional tasks, TBLT)
- determine progression through pedagogical material and suport interaction with system

⇒ Example: interactive feedback in TAGARELA

Why analyze learner language?

Learner Corpora

- offline analysis of learner language
- support effective search and analysis of annotated data
  - to provide insights into typical learner needs in FLT
  - to provide empirical evidence for SLA research, e.g.,
    - identify developmental sequences and task effects,
    - linguistic correlates of CEFR proficiency levels, or
    - native language transfer
- gold-standard training & testing data for development of NLP for learner language

Why analyze learner language?

Language Testing

- support assessment of *learner competence*
  - automate (some) grading
  - support more efficient grading by grouping learner answers
- draw valid inferences about a learner's state of knowledge
  - also central (but little discussed) for Tutoring Systems

Writer's aid tools

- provide feedback aimed at *producing text*
- identify and correct errors in orthography, grammar, usage

Why analyze native language for learners?

Input Tailoring and Input Enrichment

- identify authentic materials appropriate in readability, and
- richly representing language forms targeted by curriculum
- tailored to the needs in the learner's developmental path

⇒ Example: linguistically-aware search engine FLAIR

Input Enhancement

- Enhanced presentation of materials, adapted to learner
  - visual input enhancement supporting noticing
  - generation of annotations (e.g., vocabulary)
- Generation of exercises

⇒ Example: Input enhancement system VIEW

Multidisciplinary collaboration is required

- For NLP to address real-life needs, it must connect to
  - Second Language Acquisition (SLA) research
    - tasks, instructional interventions, relevance of input/output, interaction, meaning, focus-on-form, developmental seq.
  - Foreign Language Teaching and Learning (FLTL)
    - address teacher needs, while keeping them in charge
  - Cognitive Psychology
    - attention, memory, learning, motivation, lab studies
  - Empirical Educational Science
    - intervention studies, real-life evaluation, multi-level modeling

Interactive learning environments

- The time a student can spend with an instructor typically is very limited, limiting individual interaction and feedback.
- Individual support is increasingly essential given
  - more heterogeneous classes due to falling numbers and more children with migration background and other needs
  - informal learning environments and lifelong learning
- Intelligent tutoring systems support learners in
  - incrementally completing tasks with individual feedback
  - selecting learning materials driven by individual needs

An opportunity for CALL

- Good opportunity for developing CALL tools to
  - support acquisition of forms
  - practice language production with individual feedback
  - practice receptive skills (difficult to do in class)
  - raise linguistic awareness in general
- But existing systems typically
  - offer **limited exercise types** such as decontextualized vocabulary practice, multiple choice, point&click, form filling
  - with **feedback limited** to true/false or letter-by-letter matching of the learner response with pre-stored answers
    - Example Site: "Spanish Grammar Exercises"

Making CALL tools aware of language

- String matching as the traditional technique used to automatically analyze learner answers is effective when
  - correct answers and potential errors are predictable and listable → little well-formed or ill-formed variation
  - listable answers correspond to intended feedback
- Computational linguistic analysis must be added when
  - all possible correct and incorrect answers are not (conveniently) listable for a given activity
  - individualized feedback is desired which requires more linguistic characteristics of the learner language

An example ILTS: TAGARELA

- A concrete example for an Intelligent Tutoring System:

  TAGARELA: Teaching Aid for Grammatical Awareness, Recognition and Enhancement of Linguistic Abilities
  - intelligent web-based workbook complementing instruction
  - targeting beginning learners of Portuguese
  - designed to satisfy real-life FLT needs:
    - regular classroom instruction at OSU
    - individualized instruction at OSU
    - long-distance courses at UMass
- Focus: learner language interpretation, learner modeling, system interface design, NLP architecture, and how the system satisfies real-life needs in current FLT approaches

  (Amaral & Meurers 2006, 2008, 2009; Amaral, Meurers & Ziai 2011)

Activity types

- SLA research emphasizes the need for activities to involve both meaning and form.
- TAGARELA offers six types of activities:
  - listening comprehension
  - reading comprehension
  - picture description
  - fill-in-the-blank
  - rephrasing
  - vocabulary

= similar to traditional workbook exercises, plus audio

NLP addressing Language Learning Needs
Detmar Meurers
Introduction
Analyzing learner language
Analyzing native language
Multidisciplinarity required
Interactive learning
From CALL to ICALL
TAGARELA
Activity types
Feedback
System Architecture
Range of activity types
Content Assessment
Current projects
Adaptive materials
Readability ranking the web
FLAIR search engine
Input Enhancement
Example enhancement
Current research
Conclusion

NLP addressing Language Learning Needs
Detmar Meurers
Introduction
Analyzing learner language
Analyzing native language
Multidisciplinarity required
Interactive learning
From CALL to ICALL
TAGARELA
Activity types
Feedback
System Architecture
Range of activity types
Content Assessment
Current projects
Adaptive materials
Readability ranking the web
FLAIR search engine
Input Enhancement
Example enhancement
Current research
Conclusion

NLP addressing Language Learning Needs
Detmar Meurers
Introduction
Analyzing learner language
Analyzing native language
Multidisciplinarity required
Interactive learning
From CALL to ICALL
TAGARELA
Activity types
Feedback
System Architecture
Range of activity types
Content Assessment
Current projects
Adaptive materials
Readability ranking the web
FLAIR search engine
Input Enhancement
Example enhancement
Current research
Conclusion

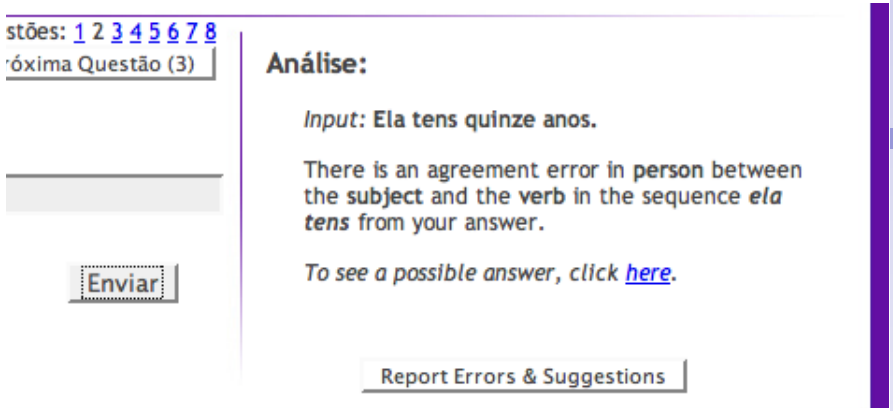
Nature of the feedback

- Which forms of feedback are most successful in fostering awareness of forms and categories?
- Some results from SLA studies on feedback carry over to human-computer interaction and CMC:
  - recasts are as effective in a dialogue system for learning English question formation (Petersen 2010)
  - recasts in synchronous CMC (Sachs & Suh 2007)
  - recasts and meta-linguistic feedback in dialogue system for maptask and appointment scheduling (Wilske 2015)

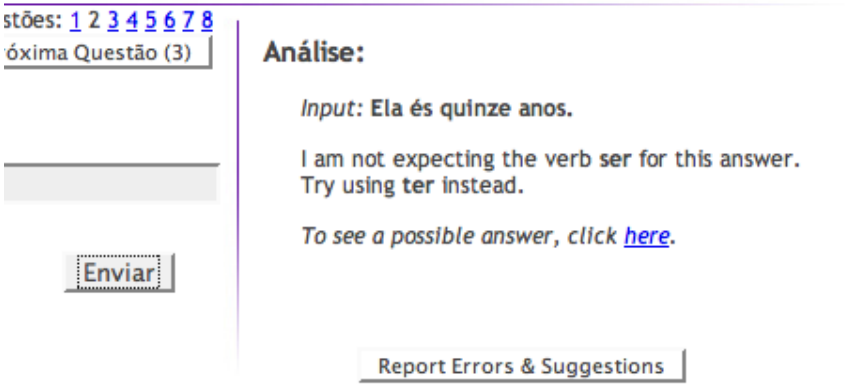
Nature of the feedback in TAGARELA

- TAGARELA provides on-the-spot feedback on
  - orthographic errors (non-word errors, spacing, capitalization, punctuation)
  - syntactic errors (nominal and verbal agreement)
  - semantic errors (missing or extra concepts, word choice)
- Nature of feedback realized for university students:
  - meta-linguistic feedback in form-focused activities
  - incidental focus-on-form in meaning-based activities
  - feedback on meaning prioritized over feedback on form

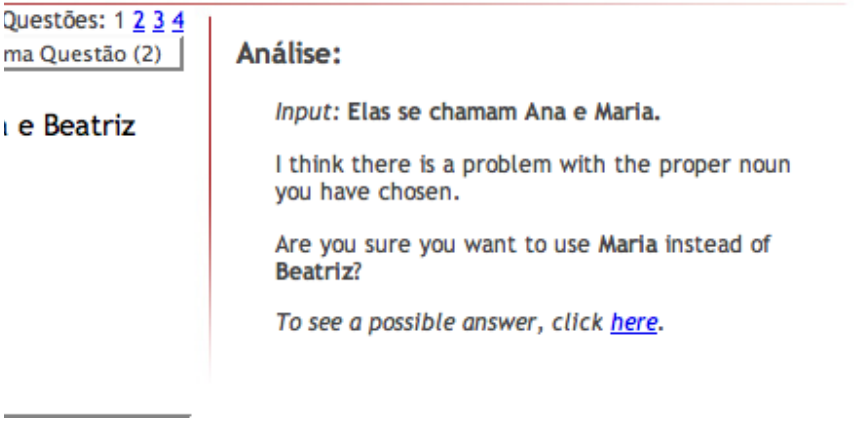
Feedback on Agreement



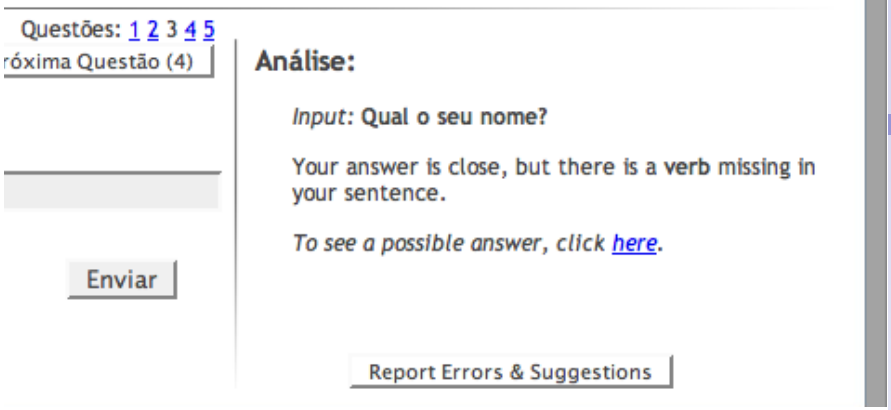
Feedback on Word Choice



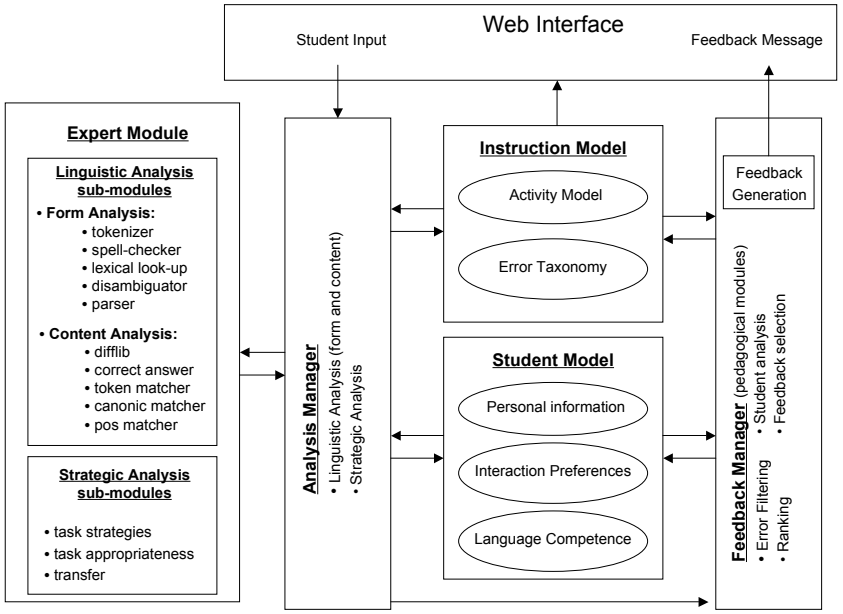
Feedback on Wrong Lexical Content



Feedback on Missing Verb



General Architecture of TAGARELA



NLP analysis modules in TAGARELA

- Form Analysis:
  - tokenizer: takes into account specifics of Portuguese (cliticization, contractions, abbreviations)
  - lexical/morphological lookup: returns multiple analyses based on CURUPIRA lexicon (Martins et al. 2006)
  - disambiguator: finite state disambiguation rules narrow down lexical information, in the spirit of Constraint Grammar (Karlsson et al. 1995; Bick 2000, 2004)
  - parser: bottom-up chart parser establishes relations to check agreement, case and global well-formedness
- Content Analysis:
  - shallow semantic matching strategies between learner answer and target answer (Bailey & Meurers 2006, 2008)
  - since 2009 topic of CoMiC project (http://purl.org/icall/comic)

Adding learner and activity models

- The TAGARELA architecture includes
  - model of domain knowledge (linguistic knowledge)
  - learner model
  - instruction/activity model
- What is the point of learner and activity models?

⇒ Providing feedback involves

- identifying properties of the learner production and
- interpreting them in terms of likely (mis)conceptions of a specfific learner trying to complete a particular activity
  - This interpretation goes beyond linguistic form as such.
  - It needs to model the learner's use of language for a specific task in a specific context (Amaral & Meurers 2008).

How to plug it all together?

- Allow the analysis manager to flexibly employ NLP modules relevant to a particular activity.
- Flexible control also relevant from NLP perspective, to support interleaving of contributions from modules, e.g.:
  - part-of-speech ambiguity in Portuguese: *a* can be a
    - preposition (*to*)
    - pronoun (*her*, clitic direct object)
    - article (*the*, feminine singular)
    - abbreviation (*association*, *alcoholic*, etc.)
  - tokenization can resolve some part-of-speech ambiguities:
    - *da* = *de* + *a* (article)
    - *vê-la* = *ver* + *a* (clitic pronoun)
    - *à* = *a* (preposition) + *a* (article)
    - *A.A.A.* = *Associação dos Alcólicos Anônimos*

→ TAGARELA tokenizer annotates some part-of-speech

Annotation-based processing

- For a flexible control structure, data structures serving as input/output of analyses need to be uniform and explicit.
- NLP analysis = enriching learner data with annotations
  - parallel to XML-based corpus annotation
- UIMA-based version of TAGARELA (Amaral et al. 2011)
  - Unstructured Information Management Architecture
- In addition to information obtained by analyzing learner production, integrate information on activity and learner.

Learner language analysis is task dependent

- In TAGARELA, different activity types require different information to interpret learner production:
  - FIB: spell-checking, lexical information
  - Rephrasing: as above + syntactic processing and basic token matcher for content assessment
  - Reading: as above + all content analysis modules
- Why not always run everything?
  - "Don't guess what you know.'
    - here: use what we know from task specification
  - The more we know the linguistic properties, the types of variation, and the potential errors the NLP needs to detect
    - the more specific information we can diagnose
    - with higher reliability

Analysis is task dependent in general

- How would you analyze the following sentences from the Hiroshima English Learners' Corpus (Miura 1998)?

  (1) *I didn't know*
  (2) *I don't know his lives.*
  (3) *I know where he lives.*

  They are taken from a translation task, for the Japanese of

  (4) *I don't know where he lives.*

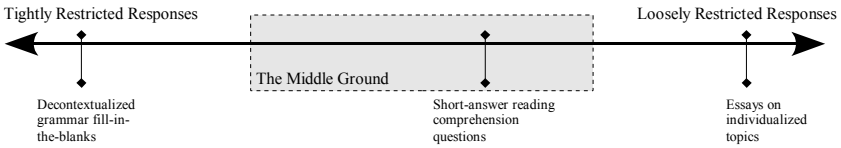
⇒ To reliably interpret learner language in ITS and learner corpus research, we should more seriously consider
  - the particular **task** and
  - the **learner** characteristics.

NLP performance confirms needs

- The best approach to grammatical error correction only reaches 39.7% precision, 30.1% recall (Ng et al. 2014)
- Inter-annotator agreement for error annotation of learner corpora is only starting to be reported (Rosen et al. 2014).
- By adding explicit
  - task design (Amaral & Meurers 2011; Quixal & Meurers 2016)
  - learner modeling (Michaud et al. 2001; Amaral & Meurers 2008)

  we can constrain the well-formed and ill-formed variation enough to obtain effective analysis of learner language.

Range of activity types



In the middle ground, there is a range of attractive activity types (reading/listening comprehension, information gap, . . .):

- a good fit with current task-based or communicative instruction settings
- effective analysis possible given predictable variability of learner responses (Quixal 2012; Quixal & Meurers 2016)
- but sophisticated meaning assessment required

  ⇒ CoMiC project: http://purl.org/icall/comic

Evaluating meaning for reading comprehension

An example from CREE (Bailey & Meurers 2008)

Q: *What are the methods of propaganda mentioned in the article?*

T: *The methods include use of labels, visual images, and beautiful or famous people promoting the idea or product. Also used is linking the product to concepts that are admired or desired and to create the impression that everyone supports the product or idea.*

Sample Learner Responses:

(5) *A number of methods of propaganda are used in the media.*
(6) *Bositive or negative labels.*
(7) *Giving positive or negative labels. Using visual images. Having a beautiful or famous person to promote. Creating the impression that everyone supports the product or idea.*

CREG Example

Q: **Was sind die Kritikpunkte, die Leute über Hamburg äußern?**

'What are the objections people have about Hamburg?'

TA: **Der Gestank von Fisch und Schiffsdiesel an den Kais .**
The stink of fish and fuel on the quays .

SemType: Der Gestank — Spelling: von Fisch — Spelling: und Schiffsdiesel — Ident: an den — Ident: Kais — Similarity: . .

SA: **Der Geruch zon Fish und Schiffsdiesel beim Hafen .**
The smell of$_{err}$ fish$_{err}$ and fuel at the port .

General CoMiC Approach

(Bailey & Meurers 2008; Meurers, Ziai, Ott & Bailey 2011a)

The overall approach has three phases:

1. **Annotation** uses NLP to enrich the student and target answers, as well as the question text, with linguistic information on different levels and types of abstraction.
2. **Alignment** maps elements of the learner answer to elements of the target response using annotation.
   - Global alignment solution computed by Traditional Marriage Algorithm (Gale & Shapley 1962)
3. **Classification** analyzes the possible alignments and labels the learner response with a binary content assessment and a detailed diagnosis code.

Types of alignment

Alignment can involve different types of representation:

Alignment Type	Example Match
Token-identical	**advertising** **advertising**
Lemma-resolved	**advertisement** **advertising**
Spelling-resolved	**campaing** **campaign**
Reference-resolved	**Clinton** **he**
Semantic similarity-resolved	**initial** **beginning**
Specialized expressions	**May 24, 2007** **5/24/2007**

Levels of alignment

Alignment can take place at different levels of representation:

Level	Example	Alignment
Tokens	**The explanation is simple.** **The reason is simple.**	**explanation** **reason**
Chunks	**A brown dog sat in a nice car.** **A nice dog sat in a car.**	**a brown dog** **a nice dog**
Depen-dency triples	**He knows the doctor.** **John knows him.**	**obj(knows, doctor)** **obj(knows, him)**

NLP tools used

Annotation Task	Language Processing Tool
Sentence Detection, Tokenization,	EN: MontyLingua (Liu 2004) DE: OpenNLP
Lemmatization	EN: PC-KIMMO (Antworth 1993) DE: TreeTagger (Schmid 1994)
Spell Checking	Edit distance (Levenshtein 1966), SCOWL word list (Atkinson 2004) igerman98 word list
Part-of-speech Tagging	TreeTagger (Schmid 1994)
Noun Phrase Chunking	CASS (Abney 1996)
Lexical Relations	WordNet (Miller 1995) GermaNet (Hamp & Feldweg 1997)
Similarity Scores	PMI-IR (Turney 2001; Mihalcea et al. 2006)
Dependency Relations	Stanford Parser (Klein & Manning 2003) MaltParser (Nivre et al. 2007)

Classification

Features

- Content Assessment is based on 13 features:

  % of Overlapping Matches:
  - keyword (head)
  - target/learner token
  - target/learner chunk
  - target/learner triple

  Nature of Matches:
  - % token matches
  - % lemma matches
  - % synonym matches
  - % similarity matches
  - % sem. type matches
  - match variety

- We combined the evidence with memory-based learning (TiMBL, Daelemans et al. 2007)
  - Trained seven classifiers using different distance metrics, overall outcome obtained through majority voting.

Meaning Assessment: Some results

- English CREE corpus: 88% accuracy of CoMiC-EN system (binary assessment, Bailey & Meurers 2008)
  - Competitive with ETS automatic scoring of native speaker short answers by C-Rater (Leacock & Chodorow 2003)
  - Alternative techniques in essay grading systems (e.g., E-Rater, Burstein et al. 2003; AutoTutor, Graesser et al. 1999) do not generalize well to short responses of 1–2 sentences.
- For German, we developed two systems
  - CoMiC-DE (Meurers, Ziai, Ott & Kopp 2011b)
  - CoSeC-DE (Hahn & Meurers 2012)

  achieving 84.6%–86.3% acccuracy on CREG corpus.
- Integration of more context information (text, question) further improves the analysis (Ziai & Meurers 2014).

Current projects

- *Linguistic form and meaning in a CL analysis of learner language. On the integration of morpho-syntactic and semantic analysis* (DFG project 2014–2019)
  - Develop NLP approach capable of interleaving bottom-up information from string with top-down information from task
  - Extend analysis of form errors (Ng et al. 2013)
- *Developing an interactive workbook for English foreign language teaching: Integrating state-of-the-art form and meaning assessment from CL into a current workbook for the Gymnasium* (DFG project 2016–2019)
  - Develop broader range of activity types integrating state-of-the-art content-assessment
  - Support and evaluate real-life use in secondary school

Teaching materials

What is offered as input to learners and how is it presented?

- Teaching materials are developed based on the contents to be communicated.
  - The complexity of the language used to express the contents so far has received only little attention.
- *Common Core State Standards* in the US raises the question of incremental textual sophistication and targets.
- In Germany, related needs are starting to be recognized:
  - "Hinführung zu Bildungssprache"
    [Progression towards academic language]
  - How can teaching materials be selected or adapted to learner populations (age, ability, migration backgr., . . .)?

Teaching materials: What can be done?

- Identify texts at level of complexity adapted to individual needs based on multi-faceted analysis of complexity
  - English (Vajjala & Meurers 2012, 2013, 2014a,b,c)
  - French (François & Fairon 2012; Todirascu, François, Gala, Fairon, Ligozat & Bernhard 2013; François & Bernhard 2014)
  - German (Hancke, Meurers & Vajjala 2012)
- Linguistically-aware search engine
  (Ott & Meurers 2010; Chinkina & Meurers 2016)
  - search for authentic texts at the right level of complexity
  - richly representing language forms targeted by curriculum
  - tailored to the needs in the learner's developmental path
- Input Enhancement of texts (Meurers et al. 2010)

Readability-ranking the web

(Vajjala & Meurers 2013)

- Are state-of-the-art readability models actually useful for classifying texts as found on the web?
  - Can we re-rank search results based on reading levels?
- Implementation details:
  - feature set inspired by SLA measures
  - WEKA linear regression, since we want output on a scale
  - trained model on 5-level WeeBit corpus
- We applied the readability model to search results obtained through BING search API.
  - took 50 search queries from a public query log
  - computed reading levels for Top-100 results

Results: Reading levels of top search results

Vajjala & Meurers (2013)

Result Rank:	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Avg.** Top100
Query:											
copyright copy law	1.8	4.6	1.4	2.7	4.6	6.2	2.7	**1.1**	3.9	5.6	4.6
halley comet	**1.7**	4.5	4.5	4.2	2.4	4.1	4.9	3.6	4.2	3.6	4.0
europe union politics	3.6	4.9	6.3	4.0	2.2	4.5	**1.5**	1.6	4.9	6.3	4.3
shakespeare	2.4	2.9	4.2	4.7	4.7	3.9	**1.5**	2.1	2.6	4.0	3.6
euclidean geometry	3.9	4.7	4.7	4.3	4.5	4.6	4.0	4.1	3.5	**2.6**	3.2
. . .											

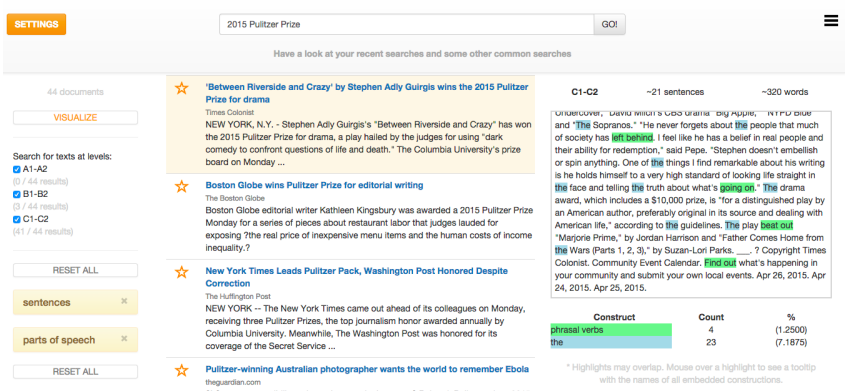
- Results:
  - avg. reading level of search results high (5 = GCSE)
  - full range of reading levels among most relevant results returned by search engine
- Re-ranking of search results potentially useful in real life

Supporting material retrieval: FLAIR

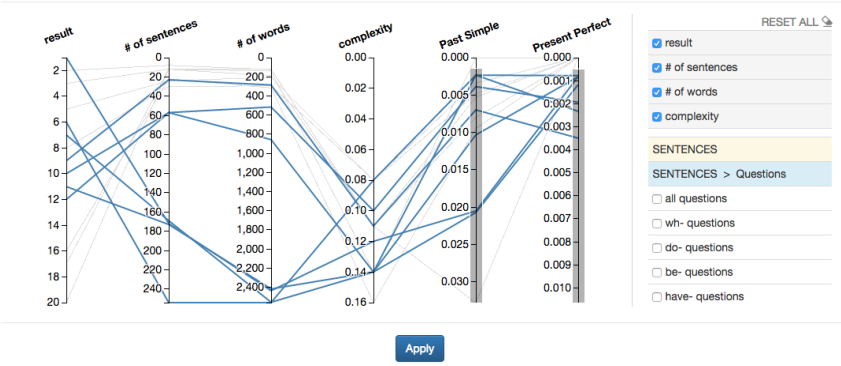
(Chinkina & Meurers 2016)

- Form-focused Linguistically Aware Information Retrieval
- identifies the 87 grammar topics spelled out in complete official English curriculum of schools (Baden-Württem.)
- designed to support teachers in identifying texts that provide the forms targeted by the curriculum
  - reranks search results based on the selected (de)prioritization of grammatical forms
  - interactively visualizes results, supporting inspection of distribution of targeted forms
- accessible at http://purl.org/icall/flair

FLAIR Interface

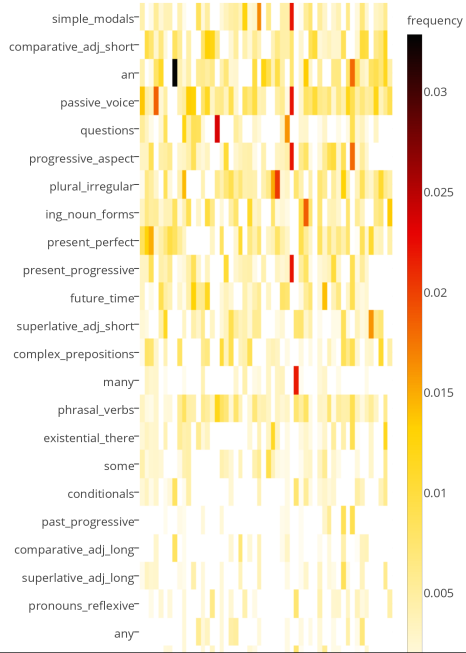


FLAIR Interactive Visualization of Results



FLAIR

Exploring the distribution of constructions in web search results



- distribution of grammatical construction across top 55 results for the query “2016 US presidential elections”
- variability shown by heat map confirms: reranking can enrich representation of many forms in curriculum
- also supports retrieving documents showcasing contrasts: *adj* vs. *adv*, *present* vs. *past simple*, etc.

Input Enhancement

- Second language learners benefit from or may require a so-called **focus on form** to overcome incomplete or incorrect knowledge (Long 1991; Lightbown 1998).
  - Focus on Form: "an occasional shift of attention to linguistic code features" (Long & Robinson 1998, p. 23).
- Strategies highlighting the salience of language forms and categories are referred to as **input enhancement** (Sharwood Smith 1993).

⇒ Automatic input enhancement for language learners

- WERTi v1 (Amaral/Meurers/Metcalf, CALICO & EUROCALL 06)
- WERTi/VIEW: Firefox Add-on + UIMA-based NLP server (Meurers et al. 2010)

WERTi system

Working with English Real Text interactively

- Provide learners of English 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)

Africa
Americas
Asia-Pacific
Europe
Middle East
South Asia
UK
England
Northern Ireland
Scotland
Wales
UK Politics
Education
Magazine
Business
Health
Science & Environment
Technology
Entertainment
Also in the news

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

News feeds

MOST POPULAR STORIES NOW

MOST SHARED | MOST READ

1 Doubt cast on Israel raid account

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

Prepositions: Practice (FIB)

Africa
Americas
Asia-Pacific
Europe
Middle East
South Asia
UK
England
Northern Ireland
Scotland
Wales
UK Politics
Education
Magazine
Business
Health
Science & Environment
Technology
Entertainment
Also in the news

E-mail this [?] 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 [?].

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

Cows moo [?] a regional twang

Listen Cow moo recordings

The farmers [?] 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 [?] 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 [?] London phonetics department

The BBC is not responsible [?] the content of external internet sites

TOP UK STORIES
- Warnings [?] Potters Bar crash
- Gaza ship Britons 'to head home'
- Walcott left out [?] England squad

News feeds

MOST POPULAR STORIES NOW

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

NLP addressing Language Learning Needs
Detmar Meurers

Introduction
- Analyzing learner language
- Analyzing native language
- Multidisciplinarity required

Interactive learning
- From CALL to ICALL
- TAGARELA
- Activity types
- Feedback
- System Architecture
- Range of activity types
- Content Assessment
- Current projects

Adaptive materials
- Readability ranking the web
- FLAIR search engine
- Input Enhancement
- Example enhancement
- Current research

Conclusion

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

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

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?

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

NLP addressing Language Learning Needs
Detmar Meurers

Introduction
- Analyzing learner language
- Analyzing native language
- Multidisciplinarity required

Interactive learning
- From CALL to ICALL
- TAGARELA
- Activity types
- Feedback
- System Architecture
- Range of activity types
- Content Assessment
- Current projects

Adaptive materials
- Readability ranking the web
- FLAIR search engine
- Input Enhancement
- Example enhancement
- Current research

Conclusion

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

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

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?

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

NLP addressing Language Learning Needs
Detmar Meurers

Introduction
- Analyzing learner language
- Analyzing native language
- Multidisciplinarity required

Interactive learning
- From CALL to ICALL
- TAGARELA
- Activity types
- Feedback
- System Architecture
- Range of activity types
- Content Assessment
- Current projects

Adaptive materials
- Readability ranking the web
- FLAIR search engine
- Input Enhancement
- Example enhancement
- Current research

Conclusion

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: 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: 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 + 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

Relation to 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)
- VIEW:
  - learner is in control of the data
  - but NLP uses 'teacher knowledge' about relevant properties to make those more prominent to the learner

Some Research Questions & Current Research

- For which language patterns is input enhanced effective?
  - Which *instances* or *aspects of their context* in a given text?
  - Using which type of input enhancement?
- Which aspects of the interaction should be tracked
  - for learners? (e.g., Open Learner Models)
  - for teachers? (e.g., satisfy grading needs)
  - for researchers? (e.g., observe incremental learning, complementing pre-/posttest design of study)
- Empirical studies needed to properly explore such issues
  - Simón Ruiz PhD project in LEAD investigates the effectiveness of input enhancement of phrasal verbs
  - A pilot study on article selection with Nicole Ziegler, Jose Luis Moreno, Wenjing Li, Simon Ruiz, Maria Chinkina, Sarah Grey, Detmar Meurers, and Patrick Rebuschat

Observe learning as it happens

Conclusion

- There is a wide range of opportunities for NLP supporting *interactive* learning environments and *adapted materials*.
- Interactive workbooks such as TAGARELA can provide immediate learner feedback on form and meaning.
  - wide range of exercise types possible using sophisticated short-answer meaning assessment → usable in real life
- Input material for the learner can be selected based on readability, curricular, and learner needs (FLAIR)
- Input enhancement tools such as WERTi/VIEW support adaptation and presentation of authentic learning material.
- Empirical studies (including tracking, pre-/posttest design) needed to validate approaches and feed back into SLA
- Interface to SLA and FLTL important
  - Special Issue of *Language Learning* to appear in 2017 targets "Language learning research at the intersection of experimental, corpus-based and computational methods"

References

Abney, S. (1996). Partial Parsing via Finite-State Cascades. In *The Robust Parsing Workshop of the European Summer School in Logic, Language and Information (ESSLLI '96)*. Prague, Czech Republic, pp. 1–8. URL http://www.vinartus.net/spa/97a.pdf.

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.

Amaral, L. & D. Meurers (2006). Where does ICALL Fit into Foreign Language Teaching? URL http://purl.org/net/icall/handouts/calico06-amaral-meurers.pdf. 23rd Annual Conference of the Computer Assisted Language Instruction Consortium (CALICO), May 19, 2006. University of Hawaii.

Amaral, L. & D. Meurers (2008). From Recording Linguistic Competence to Supporting Inferences about Language Acquisition in Context: Extending the Conceptualization of Student Models for Intelligent Computer-Assisted Language Learning. *Computer-Assisted Language Learning* 21(4), 323–338. URL http://purl.org/dm/papers/amaral-meurers-call08.html.

Amaral, L. & D. Meurers (2009). Little Things With Big Effects: On the Identification and Interpretation of Tokens for Error Diagnosis in ICALL. *CALICO Journal* 26(3), 580–591. URL http://purl.org/dm/papers/amaral-meurers-09.html.

Amaral, L. & D. Meurers (2011). On Using Intelligent Computer-Assisted Language Learning in Real-Life Foreign Language Teaching and Learning. *ReCALL* 23(1), 4–24. URL http://purl.org/dm/papers/amaral-meurers-11.html.

Amaral, L., D. Meurers & R. Ziai (2011). Analyzing Learner Language: Towards A Flexible NLP Architecture for Intelligent Language Tutors. *Computer-Assisted Language Learning* 24(1), 1–16. URL http://purl.org/dm/papers/amaral-meurers-ziai-10.html.

Antworth, E. L. (1993). Glossing Text with the PC-KIMMO Morphological Parser. *Computers and the Humanities* 26, 475–484. URL http://www.springerlink.com/content/r20w66k70976ur9l/fulltext.pdf.

Atkinson, K. (2004). Spell Checking Oriented Word Lists (SCOWL). URL http://wordlist.sourceforge.net/. Web resource.

Bailey, S. & D. Meurers (2006). Exercise-driven selection of content matching methodologies. Peer reviewed conference presentation. EUROCALL'06. September 6, 2006. University of Granada.

Bailey, S. & D. Meurers (2008). Diagnosing meaning errors in short answers to reading comprehension questions. In J. Tetreault, J. Burstein & R. D. Felice (eds.), *Proceedings of the 3rd Workshop on Innovative Use of NLP for Building Educational Applications (BEA-3) at ACL'08*. Columbus, Ohio, pp. 107–115. URL http://aclweb.org/anthology/W08-0913.

Bick, E. (2000). *The Parsing System "Palavras": Automatic Grammatical Analysis of Portuguese in a Constraint Grammar Framework*. Aarhus University Press. URL http://beta.visl.sdu.dk/~eckhard/pdf/PLP20-amilo.ps.pdf.

Bick, E. (2004). PaNoLa: Integrating Constraint Grammar and CALL. In H. Holmboe (ed.), *Nordic Language Technology, Arbog for Nordisk Sprogteknologisk Forskningsprogram 2000-2004 (Yearbook 2003)*, Copenhagen: Museum Tusculanum, pp. 183–190. URL http://beta.visl.sdu.dk/~eckhard/pdf/PaNoLa-CALL-yearbook2003.ps.pdf.

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

Burstein, J. & M. Chodorow (1999). Automated Essay Scoring for Nonnative English Speakers. In M. B. Olson (ed.), *Proceedings of a Workshop on Computer-Mediated Language Assessment and Evaluation of Natural Language Processing, Joint Symposium of the Association of Computational Linguistics (ACL-99) and the International Association of Language Learning Technologies*. pp. 68–75. URL http://www.ets.org/Media/Research/pdf/erater_acl99rev.pdf.

Burstein, J., M. Chodorow & C. Leacock (2003). Criterion: Online Essay Evaluation: An Application for Automated Evaluation of Student Essays. In *Proceedings of the Fifteenth Annual Conference on Innovative Applications of Artificial Intelligence (IAAI-03)*. Acapulco, Mexico, pp. 3–10. URL http://ftp.ets.org/pub/res/erater_iaai03_burstein.pdf.

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*. San Diego, CA.

Daelemans, W., J. Zavrel, K. van der Sloot & A. van den Bosch (2007). *TiMBL: Tilburg Memory-Based Learner Reference Guide, ILK Technical Report ILK 07-03*. Induction of Linguistic Knowledge Research Group Department of Communication and Information Sciences, Tilburg University, Tilburg, The Netherlands. URL http://ilk.uvt.nl/downloads/pub/papers/ilk.0703.pdf. Version 6.0.

Doughty, C. & J. Williams (eds.) (1998). *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press.

François, T. & D. Bernhard (eds.) (2014). *International Journal of Applied Linguistics, Special Issue on Current Research in Readability and Text Simplification*. Benjamins.

François, T. & C. Fairon (2012). An “AI readability” formula for French as a foreign language. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*.

Gale, D. & L. S. Shapley (1962). College Admissions and the Stability of Marriage. *American Mathematical Monthly* 69, 9–15. URL http://www.econ.ucsb.edu/~tedb/Courses/Ec100C/galeshapley.pdf.

Graesser, A. C., K. Wiemer-Hastings, P. Wiemer-Hastings & R. Kreuz (1999). AutoTutor: A simulation of a human tutor. *Journal of Cognitive Systems Research* 1, 35–51.

Hahn, M. & D. Meurers (2012). Evaluating the Meaning of Answers to Reading Comprehension Questions: A Semantics-Based Approach. In *Proceedings of the 7th Workshop on Innovative Use of NLP for Building Educational Applications (BEA-7) at NAACL-HLT 2012*. Montreal, pp. 94–103. URL http://purl.org/dm/papers/hahn-meurers-12.html.

Hamp, B. & H. Feldweg (1997). GermaNet – a Lexical-Semantic Net for German. In *Proceedings of ACL workshop Automatic Information Extraction and Building of Lexical Semantic Resources for NLP Applications*. Madrid. URL http://aclweb.org/anthology/W97-0802.

Hancke, J., D. Meurers & S. Vajjala (2012). Readability Classification for German using lexical, syntactic, and morphological features. In *Proceedings of the 24th International Conference on Computational Linguistics (COLING)*. Mumbay, India, pp. 1063–1080. URL http://aclweb.org/anthology-new/C/C12/C12-1065.pdf.

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.

Karlsson, F., A. Voutilainen, J. Heikkilä & A. Anttila (eds.) (1995). *Constraint Grammar: A Language-Independent System for Parsing Unrestricted Text*. No. 4 in Natural Language Processing. Berlin and New York: Mouton de Gruyter.

Klein, D. & C. D. Manning (2003). Accurate Unlexicalized Parsing. In *Proceedings of the 41st Meeting of the Association for Computational Linguistics (ACL 2003)*. Sapporo, Japan, pp. 423–430. URL http://aclweb.org/anthology/P03-1054.

Leacock, C. & M. Chodorow (2003). C-rater: Automated Scoring of Short-Answer Questions. *Computers and the Humanities* 37, 389–405.

Levenshtein, V. I. (1966). Binary Codes Capable of Correcting Deletions, Insertions, and Reversals. *Soviet Physics Doklady* 10(8), 707–710. URL http://www.mendeley.com/research/binary-codes-capable-of-correcting-insertions-and-reversals/.

Lightbown, P. M. (1998). The importance of timing in focus on form. In Doughty & Williams (1998), pp. 177–196.

Liu, H. (2004). MontyLingua: An End-to-End Natural Language Processor with Common Sense. Software Website. Media Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts. URL http://web.media.mit.edu/~hugo/montylingua/.

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. & P. Robinson (1998). Focus on form: Theory, research, and practice. In Doughty & Williams (1998), pp. 15–41.

Martins, R., R. Hasegawa & M. das Graças Nunes (2006). Curupira: a functional parser for Brazilian Portuguese. In *Computational Processing of the Portuguese*

NLP addressing Language Learning Needs

Detmar Meurers

Introduction
Analyzing learner language
Analyzing native language
Multidisciplinarity required

Interactive learning
From CALL to ICALL
TAGARELA
Activity types
Feedback
System Architecture
Range of activity types
Content Assessment
Current projects

Adaptive materials
Readability ranking the web
FLAIR search engine
Input Enhancement
Example enhancement
Current research

Conclusion

EBERHARD KARLS UNIVERSITÄT TÜBINGEN
LEAD Graduate School
69 / 69

Language, 6th International Workshop, PROPOR. Lecture Notes in Computer Science 2721. Faro, Portugal: Springer. URL http://www.springerlink.com/content/b48vjft1l88yvrj0/fulltext.pdf.

Metcalf, V. & D. Meurers (2006). Generating Web-based English Preposition Exercises from Real-World Texts. URL http://purl.org/net/icall/handouts/eurocall06-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, pp. 10–18. URL http://aclweb.org/anthology/W10-1002.pdf.

Meurers, D., R. Ziai, N. Ott & S. Bailey (2011a). Integrating Parallel Analysis Modules to Evaluate the Meaning of Answers to Reading Comprehension Questions. *IJCEELL. Special Issue on Automatic Free-text Evaluation* 21(4), 355–369. URL http://purl.org/dm/papers/meurers-ziai-ott-bailey-11.html.

Meurers, D., R. Ziai, N. Ott & J. Kopp (2011b). Evaluating Answers to Reading Comprehension Questions in Context: Results for German and the Role of Information Structure. In *Proceedings of the TextInfer 2011 Workshop on Textual Entailment*. Edinburgh: ACL, pp. 1–9. URL http://aclweb.org/anthology/W11-2401.pdf.

Michaud, L. N., K. F. McCoy & L. A. Stark (2001). Modeling the Acquisition of English: An Intelligent CALL Approach. In *Proceedings of The 8th International Conference on User Modeling*. Sonthofen, Germany, pp. 14–25. URL http://www.eecis.udel.edu/research/icicle/pubs/MiMcSt01.ps.

NLP addressing Language Learning Needs

Detmar Meurers

Introduction
Analyzing learner language
Analyzing native language
Multidisciplinarity required

Interactive learning
From CALL to ICALL
TAGARELA
Activity types
Feedback
System Architecture
Range of activity types
Content Assessment
Current projects

Adaptive materials
Readability ranking the web
FLAIR search engine
Input Enhancement
Example enhancement
Current research

Conclusion

EBERHARD KARLS UNIVERSITÄT TÜBINGEN
LEAD Graduate School
69 / 69

Mihalcea, R., C. Corley & C. Strapparava (2006). Corpus-based and Knowledge-based Measures of Text Semantic Similarity. In *Proceedings of the National Conference on Artificial Intelligence*. Menlo Park, CA: American Association for Artificial Intelligence (AAAI) Press, vol. 21(1), pp. 775–780. URL http://www.cse.unt.edu/~rada/papers/mihalcea.aaai06.pdf.

Miller, G. (1995). WordNet: a lexical database for English. *Communications of the ACM* 38(11), 39–41. URL http://aclweb.org/anthology/H94-1111.

Miura, S. (1998). Hiroshima English Learners' Corpus: English learner No. 2 (English I & English II). Department of English Language Education, Hiroshima University. http://purl.org/icall/helc.

Ng, H. T., J. Tetreault, S. M. Wu, Y. Wu & C. Hadiwinoto (eds.) (2013). *Proceedings of the Seventeenth Conference on Computational Natural Language Learning: Shared Task*. Sofia, Bulgaria: Association for Computational Linguistics. URL http://aclweb.org/anthology/W13-36.

Ng, H. T., S. M. Wu, T. Briscoe, C. Hadiwinoto, R. H. Susanto & C. Bryant (2014). The CoNLL-2014 Shared Task on Grammatical Error Correction. In *Proceedings of the Eighteenth Conference on Computational Natural Language Learning: Shared Task*. Baltimore, Maryland: Association for Computational Linguistics, pp. 1–14.

Nivre, J., J. Nilsson, J. Hall, A. Chanev, G. Eryigit, S. Kübler, S. Marinov & E. Marsi (2007). MaltParser: A Language-Independent System for Data-Driven Dependency Parsing. *Natural Language Engineering* 13(1), 1–41. URL http://w3.msi.vxu.se/~nivre/papers/nle07.pdf.

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.

NLP addressing Language Learning Needs

Detmar Meurers

Introduction
Analyzing learner language
Analyzing native language
Multidisciplinarity required

Interactive learning
From CALL to ICALL
TAGARELA
Activity types
Feedback
System Architecture
Range of activity types
Content Assessment
Current projects

Adaptive materials
Readability ranking the web
FLAIR search engine
Input Enhancement
Example enhancement
Current research

Conclusion

EBERHARD KARLS UNIVERSITÄT TÜBINGEN
LEAD Graduate School
69 / 69

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.

Quixal, M. (2012). Language Learning Tasks and Automatic Analysis of Learner Language. Connecting FLTL and NLP in the design of ICALL materials supporting effective use in real-life instruction. Ph.D. thesis, Universitat Pompeu Fabra, Barcelona and Eberhard-Karls-Universität Tübingen.

Quixal, M. & D. Meurers (2016). How can writing tasks be characterized in a way serving pedagogical goals and automatic analysis needs? *CALICO Journal* 33. URL http://purl.org/dm/papers/Quixal.Meurers-16.html. To appear.

Rosen, A., J. Hana, B. Štindlová & A. Feldman (2014). Evaluating and automating the annotation of a learner corpus. *Language Resources and Evaluation* 48(1), 65–92.

Sachs, R. & B.-R. Suh (2007). Textually enhanced recasts, learner awareness, and L2 outcomes in synchronous computer-mediated interaction. In *Oxford applied linguistics*, Oxford: Oxford University Press.

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.

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

Todirascu, A., T. François, N. Gala, C. Fairon, A.-L. Ligozat & D. Bernhard (2013). Coherence and Cohesion for the Assessment of Text Readability. In *Proceedings of the 10th International Workshop on Natural Language Processing and Cognitive Science*. URL http://www.kicorangel.com/wp-content/uploads/2013/10/NLPCS2013-proceedings.pdf.

Turney, P. (2001). Mining the Web for Synonyms: PMI-IR Versus LSA on TOEFL. In *Proceedings of the Twelfth European Conference on Machine Learning (ECML-2001)*. Freiburg, Germany, pp. 491–502.

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.

Vajjala, S. & D. Meurers (2013). On The Applicability of Readability Models to Web Texts. In *Proceedings of the Second Workshop on Predicting and Improving Text Readability for Target Reader Populations*. pp. 59–68.

Vajjala, S. & D. Meurers (2014a). Assessing the relative reading level of sentence pairs for text simplification. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*. ACL, Gothenburg, Sweden: Association for Computational Linguistics, pp. 288–297.

Vajjala, S. & D. Meurers (2014b). Exploring Measures of "Readability" for Spoken Language: Analyzing linguistic features of subtitles to identify age-specific TV programs. In *Proceedings of the Third Workshop on Predicting and Improving Text Readability for Target Reader Populations*. Gothenburg, Sweden: ACL, pp. 21–29.

NLP addressing Language Learning Needs

Detmar Meurers

Introduction
Analyzing learner language
Analyzing native language
Multidisciplinarity required

Interactive learning
From CALL to ICALL
TAGARELA
Activity types
Feedback
System Architecture
Range of activity types
Content Assessment
Current projects

Adaptive materials
Readability ranking the web
FLAIR search engine
Input Enhancement
Example enhancement
Current research

Conclusion

Vajjala, S. & D. Meurers (2014c). Readability Assessment for Text Simplification: From Analyzing Documents to Identifying Sentential Simplifications. *International Journal of Applied Linguistics, Special Issue on Current Research in Readability and Text Simplification* 165(2), 142–222.

Wiemer-Hastings, P., K. Wiemer-Hastings & A. Graesser (1999). Improving an Intelligent Tutor's Comprehension of Students with Latent Semantic Analysis. In S. Lajoie & M. Vivet (eds.), *Artificial Intelligence in Education*, IOS Press, pp. 535–542. URL http://eprints.kfupm.edu.sa/45213/1/45213.pdf.

Wilske, S. (2015). Form and meaning in dialogue-based computer-assisted language learning. Ph.D. thesis, Universität des Saarlandes, Saarbrücken. URL http://scidok.sulb.uni-saarland.de/volltexte/2015/6251/.

Ziai, R. (2009). A Flexible Annotation-Based Architecture for Intelligent Language Tutoring Systems. Master's thesis, Universität Tübingen, Seminar für Sprachwissenschaft. URL http://www.sfs.uni-tuebingen.de/~rziai/papers/Ziai-09.pdf.

Ziai, R. & D. Meurers (2014). Focus Annotation in Reading Comprehension Data. In *Proceedings of the 8th Linguistic Annotation Workshop (LAW VIII, 2014)*. COLING, Dublin, Ireland: Association for Computational Linguistics, pp. 159–168. URL http://www.aclweb.org/anthology/W14-4922.

NLP addressing Language Learning Needs

Detmar Meurers

Introduction
Analyzing learner language
Analyzing native language
Multidisciplinarity required

Interactive learning
From CALL to ICALL
TAGARELA
Activity types
Feedback
System Architecture
Range of activity types
Content Assessment
Current projects

Adaptive materials
Readability ranking the web
FLAIR search engine
Input Enhancement
Example enhancement
Current research

Conclusion