Enhancing Authentic Web Pages for Language Learners

Detmar Meurers, Ramon Ziai
Luiz Amaral, Adriane Boyd, Aleksandar Dimitrov,
Vanessa Metcalf, Niels Ott

Universität Tübingen
University of Massachusetts Amherst
The Ohio State University

Los Angeles
June 5, 2010

Our starting point

Insights from Second Language Acquisition Research

- For successful second language acquisition, meaningful, task-based use of the language to be learned is essential.
- At the same time, 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).

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

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?

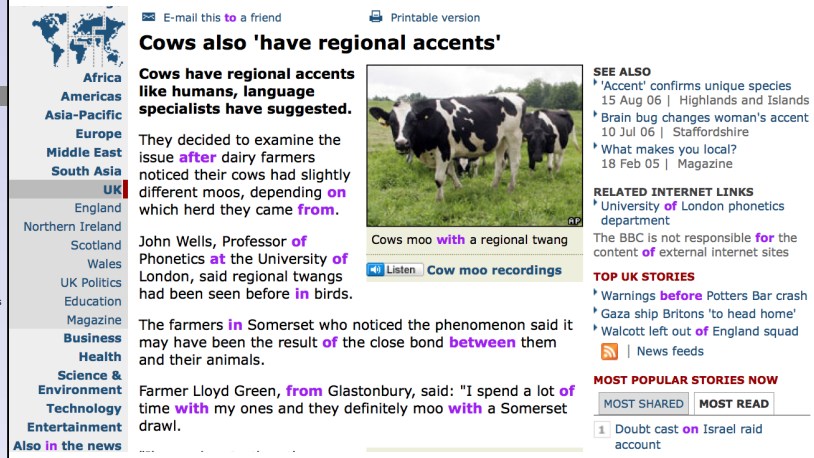
- A wide range of linguistic features can be relevant for **awareness**, incl. morphological, syntactic, semantic, and pragmatic information (Schmidt 1995).
- We focus on enhancing language patterns which are well-established difficulties for ESL learners:
  - determiner and preposition usage
  - use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

Enhancing Web Pages for Language Learners
Meurers, Ziai, Amaral, Boyd, et al.
Motivation
WERTi
What should we enhance?
How should it be enhanced?
Example activities
Prepositions
Phrasal verbs
Gerunds vs. to-infinitives
Wh-questions
Realizing WERTi
WERTi architecture
Pattern-specific NLP
Evaluation
Evaluating learning outcomes
Evaluating the NLP
Context and related work
ICALL: ILTs and ATICALL
Data-driven learning
Exercise Generation
Reading Support Tools
Summary
Outlook

Prepositions: Practice (FIB)

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

Enhancing Web Pages for Language Learners
Meurers, Ziai, Amaral, Boyd, et al.
Motivation
WERTi
What should we enhance?
How should it be enhanced?
Example activities
Prepositions
Phrasal verbs
Gerunds vs. to-infinitives
Wh-questions
Realizing WERTi
WERTi architecture
Pattern-specific NLP
Evaluation
Evaluating learning outcomes
Evaluating the NLP
Context and related work
ICALL: ILTs and ATICALL
Data-driven learning
Exercise Generation
Reading Support Tools
Summary
Outlook

Prepositions: Presentation + Interaction (Click)

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

Enhancing Web Pages for Language Learners
Meurers, Ziai, Amaral, Boyd, et al.
Motivation
WERTi
What should we enhance?
How should it be enhanced?
Example activities
Prepositions
Phrasal verbs
Gerunds vs. to-infinitives
Wh-questions
Realizing WERTi
WERTi architecture
Pattern-specific NLP
Evaluation
Evaluating learning outcomes
Evaluating the NLP
Context and related work
ICALL: ILTs and ATICALL
Data-driven learning
Exercise Generation
Reading Support Tools
Summary
Outlook

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

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.

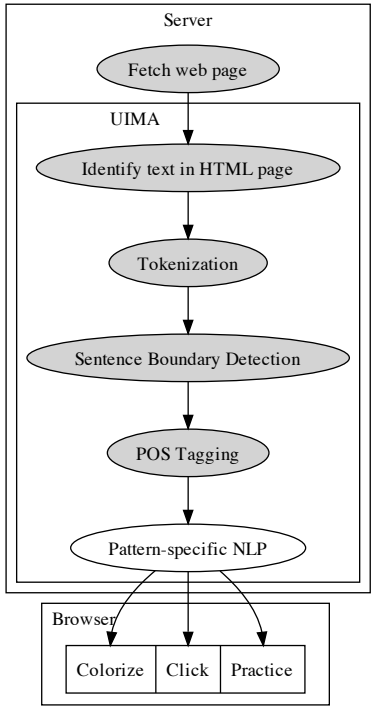
Source: http://simple.wikipedia.org/wiki/Illegal_drugs

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

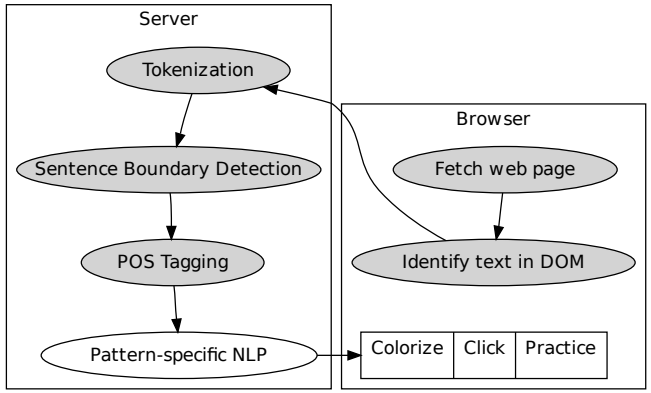
WERTi architecture

- 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.



- beta version at: `http://purl.org/icall/werti-plugin`

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. 2008b).

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 are selected by learners based on interests.
  - Next step (Ott & Meurers submitted): Develop search engine which takes into account
    - content of interest to learner
    - general readability measures (Petersen 2007; Heilman et al. 2008a; Miltsakaki & Troutt 2008)
    - language properties to be input enhanced

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?

References

Amaral, L., V. Metcalf & D. Meurers (2006). Language Awareness through Re-use of NLP Technology. Pre-conference 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-new/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-new/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.

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

Elghafari, A., D. Meurers & H. Wunsch (2010). Exploring the Data-Driven Prediction of Prepositions in English. In *Proceedings of COLING 2010*.

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.

Götz, T. & O. Suhre (2004). Design and implementation of the UIMA Common Analysis System. *IBM Systems Journal* 43(3), 476–489.

Heilman, M., K. Collins-Thompson & M. Eskenazi (2008a). An Analysis of Statistical Models and Features for Reading Difficulty Prediction. In Tetreault et al. (2008). URL http://aclweb.org/anthology-new/W08-0909.

Heilman, M., L. Zhao, J. Pino & M. Eskenazi (2008b). Retrieval of Reading Materials for Vocabulary and Reading Practice. In Tetreault et al. (2008), pp. 80–88. URL http://aclweb.org/anthology-new/W08-0910.

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.

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

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.

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. EUROCALL 2006. Granada, Spain. September 4–7, 2006.

Miltsakaki, E. & A. Troutt (2008). Real Time Web Text Classification and Analysis of Reading Difficulty. In Tetreault et al. (2008), pp. 89–97. URL http://aclweb.org/anthology-new/W08-0911.

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.

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

Ott, N. & D. Meurers (submitted). Information Retrieval for Education: Making Search Engines Language Aware URL http://purl.org/dm/papers/ott-meurers-10.html.

Petersen, S. E. (2007). Natural Language Processing Tools for Reading Level Assessment and Text Simplification for Bilingual Education. Ph.D. thesis, University of Washington. URL http://sarahpetersen.net/sarah_petersen_dissertation.pdf.

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: A tutorial on the role of attention and awareness in learning. In R. Schmidt (ed.), *Attention and awareness in foreign language learning*, Honolulu: University of Hawaii Press, pp. 1–63.

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

Tetreault, J., J. Burstein & R. De Felice (eds.) (2008). *Proceedings of the Third Workshop on Innovative Use of NLP for Building Educational Applications*. Columbus, Ohio: Association for Computational Linguistics. URL http://www.aclweb.org/anthology/W/W08/W08-09.