On Annotating Learner Language: A Computational Linguistic Perspective

Detmar Meurers
Department of Linguistics
University of Tübingen
http://purl.org/dm

Informatización de la descripción lingüística aplicada a la diagnosis experimental del aprendizaje del inglés
Scientific Meeting, Jaén, June 9, 2008

Where I'm coming from

- Use of corpora for the validation of linguistic theories (Meurers 2005, 2007; Meurers & Müller 2008)
- Automatic detection of errors in corpus annotation (Dickinson & Meurers 2003a,b, 2005a,b)
- Automatic analysis of learner language to detect
  - Word order errors (English, German) (Metcalf & Meurers 2006; Boyd & Meurers 2008)
  - Meaning errors (English) (Bailey & Meurers 2008)
- Automatic identification of Portuguese learner errors in the ICALL workbook TAGARELA (http://tagarela.osu.edu) (Amaral & Meurers 2006, 2007)

Annotated learner corpora: Starting point

- Annotated learner corpora can
  - help validate generalizations about language acquisition
  - provide a broad empirical basis for the development of new hypotheses and theories
  - inform foreign language teaching practice
- Depending on the corpus, they can support
  - qualitative and quantitative analysis
  - including longitudinal analysis
- To play these roles, the terminology used to single out the learner language aspects of interest needs to be mapped to instances in the corpus.
- Effective querying of corpora typically requires reference to annotated abstractions (linguistic classes, errors) instead of extensionally characterizing individual strings.

Some thoughts on learner language corpora

The types of learner texts included in corpora

- Most learner language corpora consist of essays.
- Yet in a typical communicative or task-based language learning setup, learners produce language in a wide range of activities, e.g.,
  - answering reading or listening comprehension questions
  - asking questions in information gap activities

⇒ To obtain corpora representative of learner language, it would be advantageous to include language produced in a variety of language activities.

- Including explicit task contexts in the meta-information of a corpus can also provide constraining information useful for interpreting learner language.
  - e.g., it's easier to infer what a learner wanted to say if one knows the text they are answering questions about.

Some thoughts on learner language corpora

The types of annotations provided

- The annotation of learner language has typically focused on errors made by the learners.
- At the same time, learner errors often are correlated with
  - specific linguistic environments (classes, constructions)?
  - specific language tasks performed by the learner (e.g., answering reading comprehension questions)?
  - or specific strategies needed to complete particular tasks (e.g., skimming, scanning)?

Some thoughts on learner language corpora

The types of annotations provided (cont.)

- Linguistic aspects of learner language are relevant for SLA research and FLT independent of errors:
  - overuse or underuse of specific language patterns.
  - measures of language development (production, understanding), e.g:
    - Developmental Sentence Scoring (Lee 1974)
    - Index of Productive Syntax (Scarborough 1990)
    - Developmental Level (Rosenberg & Abbeduto 1987; Covington et al. 2006; Lu 2008)
- Finding the relevant patterns and computing the measures requires general linguistic annotation.

  ⇒ Learner corpora should ideally provide such annotation in addition to the error annotation.

Some thoughts on learner language corpora

Ambiguity and representation

- An error annotation scheme needs to support
  - unambiguous and consistent identification of error
    - generally involves identification of target intended by learner
  - a unique representation of the identified error
- Annotation scheme design thus requires answering questions such as:
  - Where can which ambiguities be reliably resolved, given what ling. context or other information (learner, task)?
  - In a hierarchical tagset (i.e., different levels of specificity) how is consistency of level of annotation achieved?

⇒ Only distinctions reliably identified given information present in a corpus or its meta-information should be included in an annotation scheme.

Some thoughts on learner language corpora

Ambiguity and representation (cont.)

- Identifying the nature of the error
  - Example: *The man eat cheese.*
    - agreement error: *The man*$_{3s}$ *eat*$_{not(3s)}$ *cheese.*
    - tense error, intended was: *The man ate cheese.*
- Localizing and representing the error
  - Which single, unique way is chosen to *annotate* an identified error, e.g., for binary relations?
  - Example for marking a subject-verb agreement error:
    - on the subject: *The man eat cheese.*
    - on the verb: *The man eat cheese.*
    - on an annotated relation: *The man* $\rightarrow_{agr}$ *eat cheese.*
  - Problem is non-trivial given that
    - suffixes in fusioning languages combine multiple features (e.g., person, number, gender, case)
    - often multiple relations are established (e.g., D–A–A–N)

The importance of high-quality annotation

Precision of search

- By precision of search we are referring to:
  - Of the results to the query, how many represent the learner language patterns searched for?
  - False positives can result in two ways:
    - Term used for query also characterizes patterns other than the ones we are interested in.
    - Some of the annotations the query refers to are incorrect.
- Requirements on precision of search
  - for qualitative analysis: Needs to be high enough to find relevant examples among the false positives.
  - for quantitative analysis: For reliable results, very high precision is required, in particular where specific rare language phenomena are concerned (and as known from Zipf's curse, most things occur rarely).

The importance of high-quality annotation

Recall of search

- By recall of search we are referring to:
  - How many of the intended examples that in principle are in the corpus are in fact found by the query?
- Requirements on recall of search
  - for qualitative analysis: Any results found are useful, but danger of partial blindness where certain classes of examples are not captured by the query approximating target phenomenon.
  - for quantitative analysis: Maximizing recall is crucial for reliable quantitative results.

⇒ Where the query characterizing the target phenomenon is expressed in terms of the annotation, quality and consistency of the annotation is important in general, and crucial for quantitative analysis.

How to obtain high quality annotation

- Annotate corpus several times and independently, then test interannotator agreement (Brants & Skut 1998)
  - Interannotator agreement is an essential measure of whether the distinctions made in the annotation scheme can be applied consistently based on the information available in the corpus.
- Define adequate annotation scheme, with good documentation and a list of specific problematic cases, to allow for 100% agreement (Voutilainen & Järvinen 1995; Sampson & Babarczy 2003)
  - keep only distinctions which can be reliably and consistently identified and annotated uniquely
  - appendix of difficult cases and how to resolve them is essential
- Detection of annotation errors through automatic analysis of comparable data recurring in the corpus
  → DECCA project (http://decca.osu.edu)

Variation Detection for POS Annotation

Variation: material occurs multiple times in corpus with different annotations

Dickinson & Meurers (2003a) introduces the notions

- *variation nucleus*: recurring word with different annotation
- *variation n-gram*: variation nucleus with identical context

and provides an efficient algorithm to compute them.

Example: 12-gram with variation nucleus *off*

(1) *to ward off a hostile takeover attempt by two European shipping concerns*

In the two occurrences of this 12-gram in the WSJ, *off* is

- once annotated as a preposition (IN), and
- once as a particle (RP).

From Variation to Error Detection

(Dickinson & Meurers 2003a)

Variation can result from:

- *ambiguity*: different possible labels occur in different corpus occurrences
- *error*: labeling of a string is inconsistent across comparable occurrences

Heuristics used to detect POS-annotation errors:

- Variation nuclei in long *n*-grams are errors
- Distrust the fringe (= edge of variation *n*-gram)

⇒ For the WSJ, this detects 7141 non-fringe variation nuclei, with 93% error detection precision (Dickinson 2005).

Variation Detection for Syntactic Annotation

(Dickinson & Meurers 2003b)

Needed: data-driven definition of a variation nucleus as the unit of data for the comparison of syntactic annotation

But: No one-to-one mapping between word and label, as with part-of-speech.

Idea: Decompose variation nucleus detection into series of runs for all relevant string lengths, more specifically

- one-to-one mapping: string → syntactic category label (or special label NIL)
- perform runs for strings from length 1 to longest constituent in corpus

⇒ For the WSJ, this detects 6277 non-fringe variation nuclei with 71% error detection precision.

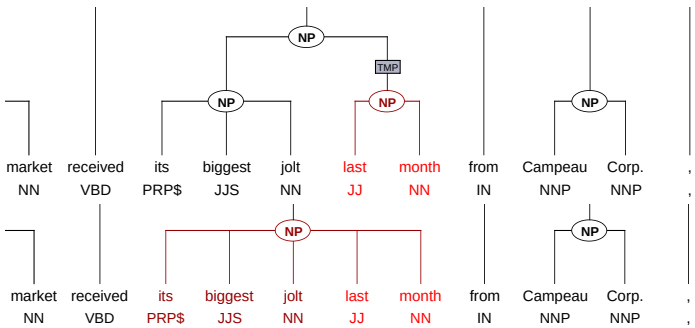
Examples from the WSJ corpus

- Variation between two syntactic category labels:

  (2) *maturity* *next Tuesday*

  labeled as **NP** twice
  PP once

- Variation between constituent and non-constituent:



Applicability of approach

- In addition to the positional and structural annotation discussed, the approach has successfully been extended to discontinuous syntactic annotation and to dependency annotation (Dickinson & Meurers 2005b; Boyd et al. 2008).
- It can in principle be applied to error annotation of learner corpora:
  - positional: one token/word
  - binary relations: lexical dependencies
  - structural domains: lists of tokens (intervals)
  - discontinuous structural domains (bags of words)
- The Python code is freely available from the project site http://decca.osu.edu.

Why is annotating word order errors relevant?

- It is hard to learn word order:
  - Language learners are known to produce a range of word order errors (cf., e.g., Odlin 1989).
  - Word order differs significantly across languages
    - → transfer errors (cf., e.g., Selinker 1972; Odlin 2003)
- Word order mistakes are common
  - e.g., Rogers (1984) reports that word order errors represent over 10% of the syntactic errors in her corpus of advanced learners of German (with L1 English).
- It is important to master word order, especially since word order errors can significantly complicate comprehension.
  - Example (Hiroshima English Learners' Corpus, HELC 1998):

    (3) *He get to cleaned his son.*

    → *He get his son to cleaned.*

Example 1: Phrasal verbs

- English learners make errors in particle placement:

  (4) a. * *so they **give up** it*

  b. * *food which will **build up** him*

  c. * *rather than **speed up** it.*

  Examples from the Chinese Learner English Corpus (CLEC 2004)
- Learners also avoid using phrasal verbs:
  - Liao & Fukuya (2002) show that Chinese learners of English avoid phrasal verbs; similar research for other L1.
  - We also found patterns of avoidance in the CLEC:
    - heavy use of pattern that is always grammatical
    - little use of patterns restricted to certain verb & object types

Example 2: Adverb placement

- Students frequently misplace adverbs

  (5) a. *they **cannot already live** without the dope.*

  b. *There **have been already** several campaigns held by 'Outdoor'.*

  c. *while any covert action **brings rarely** such negative connotations.*

  d. *It seems that the Earth **has still** a lot to reveal . . .*

  Examples from Polish part of Int. Corpus of Leaner English (PICLE 2004)
- English has many different adverbs, and the word order possibilities depend on adverb subclass distinctions.
- The rules governing adverb placement are difficult to articulate and master (and aren't strict).

Annotation of word order errors

- Error taxonomies typically focus on words to localize errors.
- Word order errors are frequently not local and depend on complex linguistic representations:
  - Grammatical function (e.g., English)
  - Topological field (e.g., German)
- Correspondingly, general learner error taxonomies typically only include a single (or no) tag for word order errors.

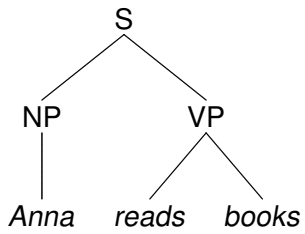
A small learner corpus (L2: German, L1: English)

- We compiled a learner corpus for
  - Development of a word order error typology
  - Evaluation of error diagnosis approaches
- Corpus of written data from introductory-level German courses at The Ohio State University:
  - Typed responses to online workbook activities
    - approx. 90,000 words in 20,000 responses; 4,000 unique responses
  - Hand-written essays from final exams
    - approx. 9,000 words in 90 essays

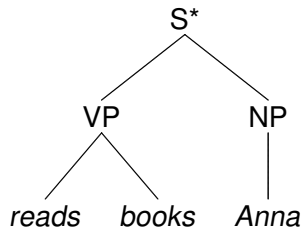
What representations do we need?

Local trees are insufficient to capture word order errors

Correct rule: S → NP VP



Mal-rule: S* → VP NP



But a *mal*-rule cannot capture **reads Anna books* (unless tree is flattened completely.)

What representations do we need?

Topological fields

- German word order is dependent on the clause type (main vs. subordinate).
- The context necessary to diagnose errors often is the entire clause.
- A topological representation of the clause can be used to make the relevant domains explicit:

Prefield	Finite/Comp Field	Middle Field	Verbal Complex	Postfield
Anna *Anna*	*liest* *reads*	*Bücher* *books*		
	daß *that*	*Anna Bücher* *Anna books*	*liest* *reads*	

Previous work on German word order errors

- Existing word order error typologies focus on:
  - Position of the finite verb relative to the clause type (main vs. subordinate) (Rogers 1984; Juozulynas 1994)
  - Edit operations relative to German topological field positions (Lüdeling et al. 2005; Hirschmann et al. 2007)
- Idea:
  - Identify word order errors in terms of the topological fields they involve,
  - but without edit operations, given that the targeted orders often are difficult to determine.

A Word Order Error Typology for German

Our proposal

1. Finite verb placement
   a) In a main clause
   b) In a subordinate clause
2. Non-finite verb placement
3. Separable prefix placement
4. Middle field
   a) Adjunct
   b) Argument
5. Noun phrase internal

Exemplifying the word order error typology

1. Finite verb placement
   a) in a main clause

(6) * *In meiner Freizeit ich bin ins Kino gegangen.*
in my free time I am into the movie gone
'In my free time I went to the movies.'

b) in a subordinate clause

(7) * *Vielleicht würde ich in München studieren, weil ich habe Familie da.*
maybe would I in Munich study since I have family there
'Maybe I would study near Munich, since I have family there.'

Exemplifying the word order error typology (cont.)

2. Non-finite verb placement

(8) * *Ich habe viel geschlafen am Wochenende.*
I have a lot slept on the weekend
'I slept a lot on the weekend.'

3. Separable prefix placement

(9) * *Ich gebe mein Geld aus Bücher und CDs.*
ich give my money out books and CDs
'I spend money on books and CDs.'

Exemplifying the word order error typology (cont.)

4. Middle field
   a) Adjunct

(5) * *Es war schwer oft.*
it was difficult often
'It was often difficult.'

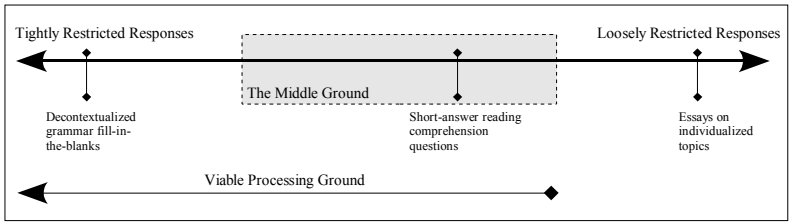
b) Argument

5. Noun phrase internal (adjective placement)

(6) * *Ich und mein Freund hatten Wetter gut.*
I and my boyfriend had weather good
'My boyfriend and I had good weather.'

Why care about meaning errors?

- Meaningful interaction in the foreign language is crucial for language learning.
- To be able to offer a wider range of activities, ICALL systems must be able to evaluate aspects of meaning.



- Loosely restricted reading comprehension (RC) questions are a good test case:
  - Common activity in real-life foreign language teaching.
  - Responses can exhibit variation on lexical, morphological, syntactic, semantic levels.
  - It is possible to specify target answers.

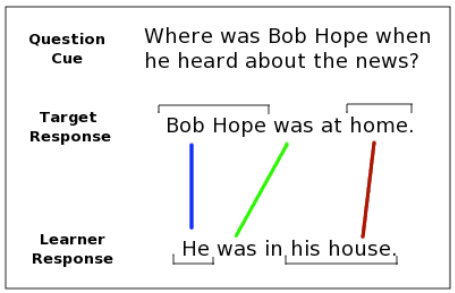
A small learner corpus (L2: English, L1: various)

- Learner corpus: 566 responses to RC questions from intermediate ESL students.
  - Development set: 311 responses from 11 students to 47 questions
  - Test set: 255 responses from 15 students to 28 questions
- Graders provided target sentences, keywords.
- Two graders annotated the data in two ways:
  - detection: Correct/Incorrect meaning
  - diagnosis (5 codes): correct; missing concept, extra concept, blend, non-answer

  Eliminated 31 responses (12%) which the graders did not agree on.
- On average, 2.7 form errors per sentence.
- Learner responses vary significantly; no full string or bag-of-word overlap with targets.

Meaning Error Detection Method

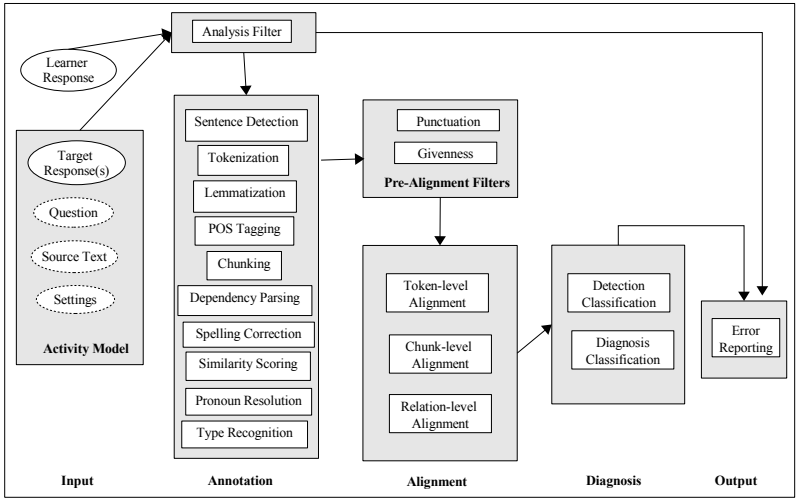
- Comparison of target and learner responses on token, chunk and relation levels, e.g.,:



Features used in Method

- 14 Features:
  - keyword/head overlap
  - target token overlap
  - learner token overlap
  - target chunk overlap
  - learner chunk overlap
  - target triple overlap
  - learner triple overlap
  - % token matches
  - % lemma matches
  - % synonym matches
  - % similarity matches
  - % sem. type matches
  - match variety
  - (sem. error detection)
- Combination of features computed by
  - manual rules
  - machine learning (TiMBL), using majority voting on available distance measures

Architecture of the Content Assessment Module



Results of meaning error detection and diagnosis

Detection	Accuracy
Baseline	50%
Development Set: Manual CAM	81%
Development Set: CAM	87%
Test Set: Manual CAM	63%
Test Set: CAM	88%

Diagnosis with 5 codes	Accuracy
Development Set	87%
Test Set	87%

- Form errors don't negatively impact results:
  - 68% of correctly diagnosed had form error
  - 53% of incorrectly diagnosed ones did so.
- No directly comparable systems, but, e.g., competitive with 85% accuracy of C-rater, automatic scoring for native speaker short answers (Leacock 2004).

On the relevance of student modeling

- Annotation of learner language (or any analysis of learner input) requires identifying the string or language pattern the learner is targeting.
- Such target identification is implicitly or explicitly based on a model of the learner: their L1, level, abilities and strategies they have demonstrated, mistakes they have made in the past, etc.
- Examples from the context of the TAGARELA system, for which we recently proposed an extension of the learner model (Amaral & Meurers to apper in 2009).

Mismatches in the identification of tokens

- Learner input: *O Amazonas fica no região norte.*
- Teacher (or ICALL system) interpretation: *no = em + o*
  - analyzed input: [$_{PP}$ em [$_{NP}$ o$_{masc}$, região$_{fem}$, norte]]

  ⇒ Agreement error between *o* and *região*.
- Student's interpretation:
  - There is no *o região norte* in the sentence I wrote.
  - I used the "preposition" *no*.

  ⇒ Wrongly interprets *no* as preposition.

Mismatches in the interpretation of accents

- Learner Input: *O vaso esta em cima de mesa.*
- Teacher (or ICALL system) interpretation:
  - The word *esta* in the learner input is a determiner.
  - There is no form of the verb *estar* in the answer.

  ⇒ The student did not include the main verb.
- Student's interpretation:
  - I included *esta* as a form of the verb *estar*. (The correct spelling is *está*.)
  - There is a verb in the sentence.

  ⇒ The lack of an accent is a spelling error.

⇒ The role of student modeling in the analysis and annotation of learner language deserves some attention.

Concluding remarks

- In this talk, I argued for
  - an emphasis on consistency of annotation
    - needs to be taken into account in annotation scheme design and documentation
    - inter-annotator agreement testing is crucial
    - automatic consistency checking can be useful
  - the annotation of learner corpora with linguistic information in addition to error annotation.
  - the creation of learner corpora stemming from a variety of tasks and activities (with explicit activity models)
- We discussed efforts to approach the annotation of
  - word order errors
  - meaning errors
- and commented on the usefulness of learner models for analyzing learner language.

References

Amaral, L. & D. Meurers (2006). Where does ICALL Fit into Foreign Language Teaching? 23rd Annual Conference of the Computer Assisted Language Instruction Consortium (CALICO), May 19, 2006. University of Hawaii.

Amaral, L. & D. Meurers (2007). Putting activity models in the driver's seat: Towards a demand-driven NLP architecture for ICALL. EUROCALL 2007, Symposium on NLP in CALL. Sept, 5-8, 2007. University of Ulster, Coleraine Campus, Ireland.

Amaral, L. & D. Meurers (to apper in 2009). 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* .

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, held at ACL 2008*. Columbus, Ohio: Association for Computational Linguistics. URL http://purl.org/dm/papers/bailey-meurers-08.html.

Boyd, A., M. Dickinson & D. Meurers (2008). On Detecting Errors in Dependency Treebanks. *Research on Language and Computation* 6(2), 113–137. URL http://purl.org/dm/papers/boyd-et-al-08.html.

Boyd, A. & D. Meurers (2008). On Diagnosing Word Order Errors. Poster presented at the CALICO Pre-Conference Workshop on Automatic Analysis of Learner Language. URL http://purl.org/net/calico-workshop-abstracts.html#6.

Brants, T. & W. Skut (1998). Automation of Treebank Annotation. In *Proceedings of New Methods in Language Processing (NeMLaP-98)*. Syndey. URL http://www.coli.uni-sb.de/~thorsten/publications/Brants-Skut-NeMLaP98.ps.gz.

CLEC (2004). Chinese Learner English Corpus. Web interface to Corpus.

Covington, M. A., C. He, C. Brown, L. Naci & J. Brown (2006). *How complex is that sentence? A proposed revision of the Rosenberg and Abbeduto D-Level Scale*. CASPR Research Report 2006-01, The University of Georgia, Artificial Intelligence Center, Athens, GA.

Dickinson, M. (2005). Error detection and correction in annotated corpora. Ph.D. thesis, The Ohio State University.

Dickinson, M. & W. D. Meurers (2003a). Detecting Errors in Part-of-Speech Annotation. In *Proceedings of the 10th Conference of the European Chapter of the Association for Computational Linguistics (EACL-03)*. Budapest, Hungary, pp. 107–114. URL http://ling.osu.edu/~dm/papers/dickinson-meurers-03.html.

Dickinson, M. & W. D. Meurers (2003b). Detecting Inconsistencies in Treebanks. In *Proceedings of the Second Workshop on Treebanks and Linguistic Theories (TLT-03)*. Växjö, Sweden, pp. 45–56. URL http://purl.org/dm/papers/dickinson-meurers-tlt03.html.

Dickinson, M. & W. D. Meurers (2005a). Detecting Annotation Errors in Spoken Language Corpora. In *The Special Session on treebanks for spoken language and discourse at NODALIDA-05*. Joensuu, Finland. URL http://purl.org/~dm/papers/dickinson-meurers-nodalida05.html.

Dickinson, M. & W. D. Meurers (2005b). Detecting Errors in Discontinuous Structural Annotation. In *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL-05)*. pp. 322–329. URL http://aclweb.org/anthology/P05-1040.

HELC (1998). Hiroshima English Learners' Corpus. Data available on webpage. URL http://home.hiroshima-u.ac.jp/d052121/eigo2.html.

Hirschmann, H., S. Doolittle & A. Lüdeling (2007). Syntactic Annotation of Non-Canonical Linguistic Structures. In *Proceedings of Corpus Linguistics 2007*.

Juozulynas, V. (1994). Errors in the Compositions of Second-Year German Students: An Empirical Study for Parser-Based ICALI. *CALICO* 12(1), 5–17.

Leacock, C. (2004). Scoring Free-Responses Automatically: A Case Study of a Large-Scale Assessment. *Examens* 1(3).

Lee, L. (1974). *Developmental Sentence Analysis*. Evanston, IL: Northwestern University Press.

Liao, Y. D. & Y. J. Fukuya (2002). Avoidance of phrasal verbs: The case of Chinese learners of English. *Second Language Studies* 20(2), 71–106.

Lu, X. (2008). Automatic measurement of syntactic complexity using the revised developmental scale. In *Proceedings of the 21st International Florida Artificial Intelligence Research Society Conference (FLAIRS-08)*. Coconut Grove, FL: AAAI Press.

Lüdeling, A., M. Walter, E. Kroymann & P. Adolphs (2005). Multi-level error annotation in learner corpora. In *Proceedings of Corpus Linguistics*. Birmingham.

Metcalf, V. & D. Meurers (2006). When to Use Deep Processing and When Not To – The Example of Word Order Errors. Pre-conference Workshop on NLP in CALL – Computational and Linguistic Challenges. CALICO 2006. May 17, 2006. University of Hawaii.

Meurers, D. & S. Müller (2008). Corpora and Syntax (Article 44). In A. Lüdeling & M. Kytö (eds.), *Corpus Linguistics. An International Handbook*, Berlin: Mouton de Gruyter, Handbooks of Linguistics and Communication Science. URL http://purl.org/dm/papers/meurers-mueller-07.html.

Meurers, W. D. (2005). On the use of electronic corpora for theoretical linguistics. Case studies from the syntax of German. *Lingua* 115(11), 1619–1639. URL http://purl.org/dm/papers/meurers-03.html.

Meurers, W. D. (2007). Advancing Linguistics Between the Extremes: Some thoughts on Geoffrey Sampson's *Grammar without Grammaticality*. *Corpus Linguistics and Linguistic Theory* 3(1), 49–55. URL http://purl.org/net/dm/papers/meurers-cllt07.html.

Odlin, T. (1989). *Language Transfer: Cross-linguistic influence in language learning*. New York: Cambridge University Press.

Odlin, T. (2003). Cross-linguistic Influence. In C. Doughty & M. Long (eds.), *Handbook on Second Language Acquisition*, Oxford: Blackwell, pp. 436–486.

PICLE (2004). Polish portion of the International Corpus of Learner English. Web interface to Corpus. URL http://elex.amu.edu.pl/~przemka/concord2advr/search_adv_new.html.

Rogers, M. (1984). On Major Types of Written Error in Advanced Students of German. *International Review of Applied Linguistics in Language Teaching* XXII(1).

Rosenberg, S. & L. Abbeduto (1987). Indicators of linguistic competence in the peer group conversational behavior of mildly retarded adults. *Applied Psycholinguistics* 8, 19–32.

Sampson, G. & A. Babarczy (2003). Limits to annotation precision. In *Proceedings of the 4th International Workshop on Linguistically Interpreted Corpora (LINC-03)*. pp. 61–68. URL http://www.grsampson.net/Alta.html.

On Annotating Learner Language: A CL Perspective

Detmar Meurers

Introduction

On learner corpora
Useful text types
Useful types of annotation
Ambiguity and representation

Annotation quality
Why it's important
How to obtain high quality
Variation Detection for POS
Variation → Error Detection
Syntactic Annotation
Applicability

Word order errors
Why are they relevant?
Examples
German Corpus
Word order error typology

Meaning errors
Why are they relevant?
English Corpus
Method
Results

Student modeling

Conclusion

Eberhard Karls Universität Tübingen

38 / 38

Scarborough, H. S. (1990). Index of Productive Syntax. *Applied Psycholinguistics* 11(1), 1–22.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics* 10(3), 209–231.

Voutilainen, A. & T. Järvinen (1995). Specifying a shallow grammatical representation for parsing purposes. In *Proceedings of the 7th Conference of the EACL*. Dublin, Ireland. URL http://www.aclweb.org/anthology-new/E95-1029.

On Annotating Learner Language: A CL Perspective

Detmar Meurers

Introduction

On learner corpora
Useful text types
Useful types of annotation
Ambiguity and representation

Annotation quality
Why it's important
How to obtain high quality
Variation Detection for POS
Variation → Error Detection
Syntactic Annotation
Applicability

Word order errors
Why are they relevant?
Examples
German Corpus
Word order error typology

Meaning errors
Why are they relevant?
English Corpus
Method
Results

Student modeling

Conclusion

Eberhard Karls Universität Tübingen

38 / 38