

# Error-tagging the German component of LINDSEI: Methodology and application

MIMS 2010, Hamburg, 06/10 - 08/10/2010

Susanne Kämmerer

Sandra Götz

Justus Liebig University Giessen, Germany

Macquarie University Sydney, Australia

# Overview

---

1. LINDSEI
2. The Louvain error-tagging system
3. Adaptations
4. Pilot Study
5. Outlook

# 1. LINDSEI

---

- Louvain International Database of Spoken English Interlanguage (cf. Gilquin et al. forthcoming)
- initiated: Université Catholique de Louvain, German subcorpus: Justus Liebig University Giessen
- spoken learner language
- 11 complete national subcorpora (11 L1-backgrounds: Bulgarian, Chinese, Dutch, French, German, Greek, Italian, Japanese, Polish, Spanish, Swedish)  $\beta$  still expanding
- 50 interviews (monologic / dialogic parts)
- each interview ~ 2000 words
- $\beta$  only LINDSEI-GE: fully error-tagged
- $\beta$  ~ 1300 errors in LINDSEI-GE (errors analysed according to IF-status)

## 2. The Louvain error-tagging system (1)

---

- ⌋ German subcorpus of LINDSEI was error-tagged using the *UCL Error-Tagging Manual Version 1.2* (Dagneaux et al. 2005)
- ⌋ *Centre of English Corpus Linguistics, Université catholique de Louvain, Belgium*
- ⌋ Louvain error-tagging system and its accompanying manual were specifically designed for ICLE

## 2. The Louvain error-tagging system (2)

---

- flat level annotation system
- maximum of four levels: category classification followed by up to three modifications refining the error, of which the fourth level might be added if the researchers wish to specify that transfer might be at work
- follows the principles for creation of an error-annotation system that is consistent, informative, flexible and reusable (Granger 2003:467)

Level 1	Level 2	Level 3	[Level 4]
<u>Grammar</u> : (G)	<u>Determiners</u> : (GD)	<u>Demonstrative determiners</u> (GDD) <u>Possessive determiners</u> (GDO) <u>Indefinite determiners</u> (GDI) <u>Determiner other</u> (GDT)	
	<u>Articles</u> (GA)		
	<u>Nouns</u> (GN)	Noun <u>c</u> ase (GNC) Noun <u>n</u> umber (GNN)	
	<u>Pronouns</u> (GP)	<u>Demonstrative pronouns</u> (GPD) <u>Personal pronouns</u> (GPP) <u>Possessive pronouns</u> (GPO) <u>Indefinite pronouns</u> (GPI) <u>Reflexive and reciprocal pronouns</u> (GPF) <u>Relative and interrogative pronouns</u> (GPR) <u>Unclear pronominal reference</u> (GPU)	
	<u>Adjectives</u> (GADJ)	Adjective <u>o</u> rders (GADJO) Adjective <u>n</u> umber (GADJN) <u>Comparative / Superlative</u> (GADJCS)	
	<u>Adverbs</u> (GADV)	Adverb <u>o</u> rders (GADVO)	
	<u>Verbs</u> (GV)	Verb <u>n</u> umber (GVN) Verb <u>m</u> orphology (GVM) <u>Non-Finite / Finite verb forms</u> (GVNF) Verb <u>v</u> oice (GVV) Verb <u>t</u> ense (GVT) <u>Auxiliaries</u> (GVAUX)	
	<u>Word Class</u> (GWC)		

## 2. The Louvain error-tagging system (3)

---

### 56 error tags within seven main error categories

- ⌋ [Form (F)]
- ⌋ Lexicogrammar (X)
- ⌋ Lexis (L)
- ⌋ Style (S)
- ⌋ Grammar (G)
- ⌋ Words (redundant / missing) / Word order (W)
- ⌋ [Punctuation (Q)]
- +
- ⌋ Infelicities (Z) (distinguished from errors)

<b>Error category</b>	<b>Error tag</b>	<b>Corpus example</b>
<p><b>Form:</b> This category groups together all words that are non-existent in English and other errors which are obviously of a formal nature.</p> <p>Note: Subcategory Spelling does not apply to speech.</p>	F	(1) <B> you have to do two (FM) practicum\$ \$periods of teaching practice\$ <\B> (GE009)
<p><b>Lexicogrammar:</b> This category consists of errors where the morpho-syntactic properties of a word have been violated. [ý ] The most important areas of error are countable/uncountable nouns, non-finite / finite complementation of nouns, verbs, adjectives and verbs and dependent prepositions.</p>	X	(2) <B> there was something erm . I was quite .. (XADJPR) disappointed of \$disappointed with\$ <\B> (GE013)
<p><b>Lexis:</b> This general category deals with errors involving the semantic (conceptual, collocational, or connotative) properties of words or phrases. It is divided into three subcategories: Lexical Single, Lexical Phrase and Connectors.</p>	L	(3) <B> she probably asks him to (LS) make \$do\$ a really nice portrait of her <\B> (GE009)
<p><b>Style:</b> This category is exclusively restricted to incomplete sentences [ý ] and unclear sentences [ý ].</p>	S	(4) (SI) Another example \$Another example is: \$ Yesterday we spoke about the GULF War (ý ).



<b>Error category</b>	<b>Error tag</b>	<b>Corpus example</b>
<p><b>Grammar:</b> This major category groups together errors that break general rules of English grammar.</p>	G	(5) <B> or they (GVT) have been \$were\$ there forty years ago <\B> (GE009)
<p><b>Word redundancy / Word missing / Word order:</b> W[ord] R[edundant] involves unnecessary repetitions of words. However, not all cases where a word is redundant should be tagged WR. W[ord] M[issing]: This sub-category for errors involving the omission of words. W[ord] O[rder]: is for problems of word order that do not fall into the categories of Adverb Order (GADVO) [ý ] or Adjective Order (GADJO).</p>	W	(6) <B> I ha= had to listen to jazz . from from the cradle . (WRS) on \$0\$ <laughs> probably to the grave <\B> (GE013)
<p><b>Punctuation:</b>  Note: Category does not apply to speech.</p>	Q	
<p><b>Infelicities:</b> [ý ] [T]he tagging manual makes a crucially important distinction between real errors, on the one hand, and infelicities, on the other. Infelicities are tagged Z and include the following problems: register problems, questions of political correctness and stylistic problems.</p>	Z	(7) <B> quite often we (Z) had the possibility \$were able\$ to .. erm . to sleep in a barn <\B> (GE023)

## 2. The Louvain error-tagging system (4)

---

Error categories not accounted for:

- Ⓟ Pronunciation errors
- Ⓟ Inconsistency with use of variety
- Ⓟ Self-correction
- Ⓟ Error-Repetition
  
- Ⓟ Learner-centred identification and explanation of errors
  
- Ⓟ Double Tag
- Ⓟ Treble Tag

# 3. Adaptation: Pronunciation

---

- ⌋ **Trigger:** simple stress shift, mispronunciation of sounds specific to a particular language, rhotic / non-rhotic pronunciation, etc.
- ⌋ **Example:** (8) \ˈblak-ˌbʊrd\ (compound) vs. \ˌblak ˈbʊrd\ (phrase)
- ⌋ **Suggested tag:** (P) (Pronunciation) as a level-1 error category with no indication of the type of mispronunciation. The researcher may want to consult the sound files in order to listen to the erroneous form and find out about the type of mispronunciation
- ⌋ **Problem:** rendering the annotator's job difficult
- ⌋ **Solution:** not all errors should be tagged, but only those mispronounced forms that result in change of word class or meaning change
- ⌋ **Current status:** not yet included

### 3. Adaptation: Inconsistency with use of variety

---

- ⌋ **Trigger:** use of prepositions and infinitive/gerund sometimes depends on variety
- ⌋ **Example:** (9) <B> I wasn't really good in English in school <\B> (GE009)
- ⌋ **Suggested tag:** (Y) (**Variety**) The tag is to be treated as a sub-level error category with the correction given directly after the identified erroneous sequence

(10) <B> I wasn't really good (LSPY) in \$at\$ English (LSPY) in \$at\$ school <\B>

- ⌋ **Problem 1:** rendering the annotator's job tremendously difficult
- ⌋ **Problem 2:** most of the students do not speak a particular style but use a more transatlantic mix. The UCL Error-Tagging Manual suggests tagging such instances (Z) (Infelicities) since it would be too descriptive a position
- ⌋ **Solution:** leave out?
- ⌋ **Current status:** tagged as (Z)

# 3. Adaptation: Self-correction

---

- ⌋ **Trigger:** marks awareness of an error
- ⌋ **Example:** (11) <B> I am . er was in Rome last year <\B> (GE044)
- ⌋ **Suggested Tag:** (S) (Self-correction). The tag is to be treated as a sub-level error category with no correction given directly after the erroneous sequence since *self-correction* is not regarded as having produced an erroneous form but rather shows awareness of the same:

(12) <B> I (GVTS) am . er was in Rome last year <\B>

- ⌋ **Problem:** -
- ⌋ **Solution:** -
- ⌋ **Current status:** not yet included

# 3. Adaptation: Error-repetition

---

- ⌋ **Trigger:** strictly speaking they represent only one error but what we get is two erroneous lexical items in a row
- ⌋ **Example:** (13) <B> haven't been there and . or they have been there forty years ago <\B> (GE009)
- ⌋ **Suggested Tag:** (R) (Repetition). The suggested tag is to be treated as a sub-level error category with the correction given directly after the erroneous sequence. The first occurrence of the error is tagged as suggested by the UCL Error-Tagging Manual. The second occurrence is to be tagged using the new tag:

(14) <B> (GVT) haven't been \$weren't\$ there and . or they (GVTR) have been \$were\$ there forty years ago <\B>

- ⌋ **Problem:** -
- ⌋ **Solution:** -
- ⌋ **Current status:** not yet included

### 3. Adaptation: Learner-centred identification and explanation of errors (1)

---

- ⌋ **Trigger:** evaluation of learner language mainly follows two principles:
  - ⌋ (a) the assumptions about a learner's competence in the target language are based on their performance, and
  - ⌋ (b) the assumptions about the causes of errors are based on the researcher's opinion alone
  
- ⌋ **Question:** performance mirrors competence in the case of spontaneous speech production in the target language?
  
- ⌋ **Hypothesis:** not all errors found in learner output are genuine competence deficits (disturbance phenomena) + the inclusion of the learner-centred explanation of the sources of errors reduces the subjectivity of the researcher to a minimum (Dulay et al. 1982:144, Dagneaux et al. 1996:166, James 1998:102)

### 3. Adaptation: Learner-centred identification and explanation of errors (2)

---

- ⌋ **Suggested tag:** (ID) (IDentification). The suggested tag is to be treated as a sub-level error category with the correction given directly after the identified erroneous sequence

(15) <B> it worked actually very (GWCID) good \$well\$ and I <\B>

- ⌋ **Problem:** learners' willingness to participate
- ⌋ **Solution:** making participation compulsory
- ⌋ **Current status:** included in corpus for pilot study (see Kämmerer 2009, Kämmerer forthcoming), but not in LINDSEI-GE



# 3. Adaptation: Interference-induced errors

---

- ⌋ **Trigger:** not all transfer-induced errors constitute formally similar words in the foreign language and in the learner's L1
- ⌋ **Example:** (16) He (LSF) pretends \$claims\$ to be ill (F.: *prétendre*) (Louvain error-tagging manual)
- ⌋ **Suggested Tag:** (IF) (Interference). The suggested tag is to be treated as a sub-level error category with the correction given directly after the identified erroneous sequence

(17) <B> trying to find (WMIF) 0 \$our way\$ back to the hotel <\B>

- ⌋ **Problem:** subjective/need for competent speaker of learner's L1
- ⌋ **Solution:** learner's help as regards explanation for cause of errors
- ⌋ **Current status:** included in LINDSEI-GE + complete list of all ~1300 errors (including IF-status and source)

### 3. Adaptation: Double Tags (1)

---

- ⌋ **Trigger:** UCL Error-Tagging Manual states that whenever there are two types of errors in the same word or phrase, then this is to be double-tagged:

(20) GE untagged (GE018: LPF + GVT):

<B> but h= . he actually only drew her how she really looks like and she is not satisfied <\B>

(21) Error-tagged (according to the UCL Error-Tagging Manual):

<B> but h= . he actually only drew her (LPF) how she really looks like \$the way she really looked\$ (GVT) how she really looks like \$the way she really looked\$ and she is not satisfied <\B>

- ⌋ **Problem with this kind of tagging:** not obvious from the error-tagged corpus whether what the researcher is dealing with is (a) a repetition with one error each or (b) a double tag

## 3. Adaptation: Double Tags (2)

---

- ⌋ **Suggested Tag:** (DTG) (Double Tag). Where there are two types of errors in the same word / phrase, it is double-tagged. In order to indicate that there are two errors in one word / phrase rather than the speaker's repetition of an erroneous word / phrase or two different types of errors in two independent words, the tag (DTG) is placed immediately in front of the first error tag

(22) <B> but h= . he actually only drew her (DTG) (LPF) how she really looks like \$the way she really looked\$ (GVT) how she really looks like \$the way she really looked\$ and she is not satisfied <\B>

- ⌋ **Problems:** -
- ⌋ **Solutions:** -
- ⌋ **Current status:** included in LINDSEI-GE

# 3. Adaptation: Treble Tags

---

- ⌋ **Trigger:** whenever there are three types of errors in the same word or phrase, then this is to be treble-tagged accordingly:

(23) GE untagged (GE009: WM + GNN + WO):

<B> and erm I stayed eight month in Wales <\B>

(24) Error-tagged (according to the UCL Error-Tagging Manual):

<B> and erm I stayed (WM) eight month \$for eight months\$ eight (GNN) month \$eight months\$ (WO) eight month in Wales \$in Wales for eight months\$ <\B>

- ⌋ **Suggested Tag:** (TTG) (Treble Tag). Where there are three types of errors in the same word / phrase, a treble tag is used. The tag (TTG) is placed immediately in front of the first error tag.

(25) <B> and erm I stayed (TTG) (WM) eight month \$for eight months\$ eight (GNN) month \$eight months\$ (WO) eight months in Wales \$in Wales for eight months\$ <\B>

- ⌋ **Problems:** -
- ⌋ **Solutions:** -
- ⌋ **Current status:** included in LINDSEI-GE

## 4. Pilot study: Accuracy in advanced learner language

---

Advanced learners' output is said to be highly idiosyncratic (cf. Cobb 2003)

→ Are there any patterns in advanced learners' errors?

**Database:** error-tagged version of LINDSEI-GE

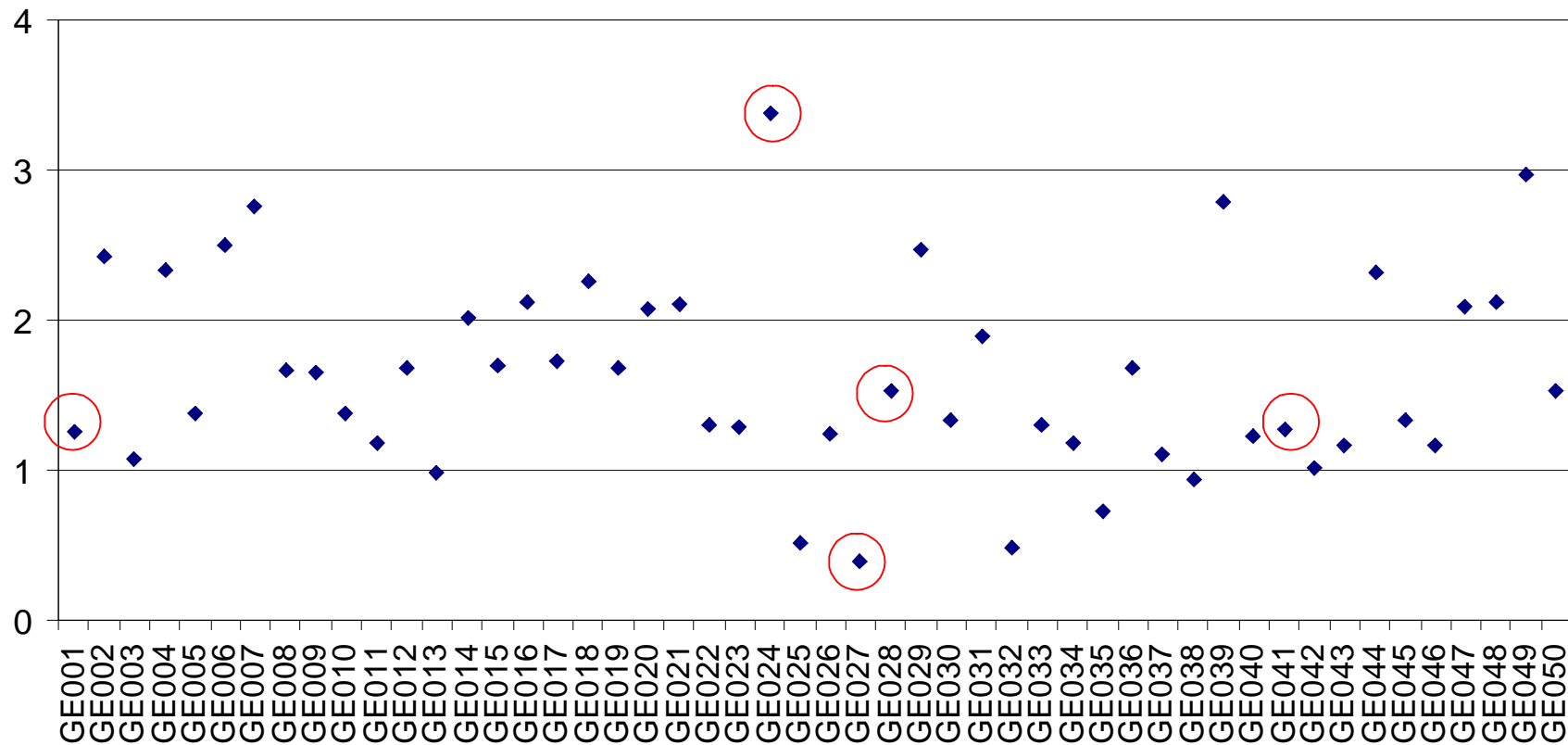
**Step 1:** extraction of number of errors quantitatively

**Step 2:** selection of 5 learners for qualitative error analysis

**Step 3:** test correlation with native speaker ratings of overall oral proficiency

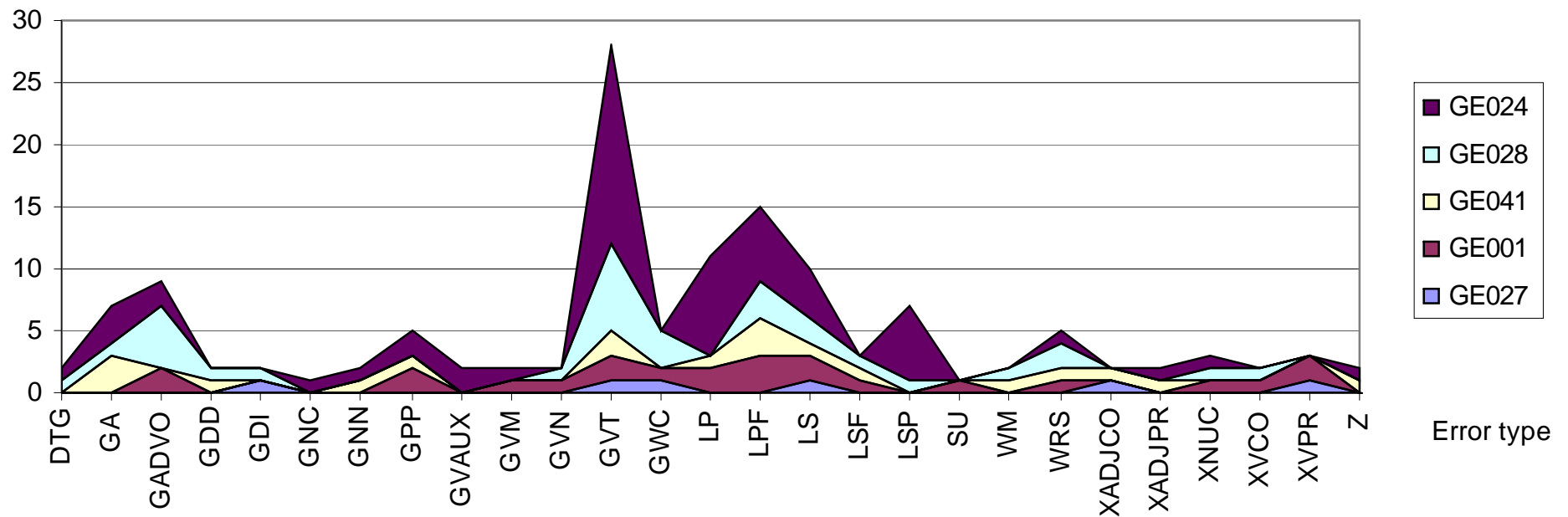
# 4.1 Quantitative error analysis

Figure 1: Errors per one hundred words in LINDSEI-GE (cf. Brand & Götz forthcoming)



# 4.2 Qualitative error analysis

Figure 2: Error types of 5 learners of LINDSEI-GE (cf. Brand & Götz forthcoming)



## 4.2 Qualitative error analysis (1)

---

### **GVT (20%) = Grammar/Verbs/Tense**

(26) We didn't know where the campus was and the bus (GVT) is . is driving \$was driving\$ through (GE024)

(27) I actually took out a map and [begin laughter] (GVT) was looking \$looked\$ where I was (GE001)

### **LPF (11%) = Lexical Phrase/False Friend**

(28) I (LPF) had the luck \$was lucky\$ to have a vacation (GE041)

(29) We pay fifty Euros (LPF) in the semester \$per semester\$ (GE028)



## 4.2 Qualitative error analysis (2)

---

### **LP (8%) = Lexical Phrase**

(30) It is always easy for me to (LP) get into contact \$make contact\$  
(GE024)

(31) I was sitting there and . (LP) looking into \$staring at\$ this TV (GE001)

### **LS (6%) = Lexical Single**

(32) But it was erm . erm . a (LS) parody \$caricature\$ . of me (GE027)

(33) I don't know . getting money, (LS) doing \$having\$ a career (GE001)

### **GADVO (6%) = Grammar/Adverb/Order**

(34) We could travel (GADVO) a little around \$around a little\$ (GE028)

(35) (GADVO) alone the weather \$the weather alone\$ (GE024)

## 4.3 Native speaker ratings (1)

---

- Anonymus online-survey with *QuestionPro*
- Raters are staff and PhD-students from the faculties of Arts and the Human Sciences from Macquarie University Sydney (N=50)
- Raters listened to these five learners and rated their overall oral proficiency from 1-10

## 4.3 Native speaker ratings (2)

Table 1: NS-ratings of overall proficiency compared to fluency and errors phw  
(N=50, ICC=0.89)

Learner	Errors phw	Overall Proficiency
Learner GE024	3.38	7.68
Learner GE027	0.40	6.80
Learner GE028	1.53	7.96
Learner GE041	1.28	7.10
Learner GE001	1.26	7.70
Correlation	r=0.560, p>0.05	

# 5. Outlook

---

- Ⓟ Error tagging of a learner corpus can be very beneficial for learner language analysis
  
- Ⓟ Pilot study has revealed that there are clear patterns across learners concerning the error types
  - more emphasis on these areas in FLT at the university level
  
  - No obvious trend for a correlation with NS-ratings
    - detailed qualitative analysis of all learners and other variables is needed
  
- Ⓟ Match results with other national (error-tagged) subcorpora of LINDSEI

# References

- Brand, C. & S. Götz (forthcoming): "Fluency and accuracy in advanced spoken learner language: a multi-method approach". *International Journal of Corpus Linguistics* 16 (2).
- Brand, C. & S. Kämmerer (2006): *The Louvain International Database of Spoken English Interlanguage (LINDSEI): Compiling the German Component*. *Corpus Technology and Language Pedagogy: New Resources, New Tools, New Methods*, eds. S. Braun, K. Kohn & J. Mukherjee. Frankfurt am Main: Peter Lang. 127-140.
- Cobb, T. (2003): "Analyzing late interlanguage with learner corpora: Quebec replications of three European studies", *Canadian Modern Language Review* 59 (3), 393-423.
- Corder, S. P. (1974): "Error Analysis". In: J. P. B. Allen and S. Pit Corder (eds.) *Techniques in Applied Linguistics*. London: Oxford University Press, pp. 122-154
- Dagneaux, E. et al. (1996): "Computer-aided error analysis", *System* 26, 163-174.
- Dagneaux, E., S. Denness, S. Granger, F. Meunier, J. Neff & J. Thewissen (2005): *UCL Error-Tagging Manual Version 1.2*. Centre for English Corpus Linguistics Université Catholique de Louvain: Louvain-la-Neuve. Unpublished manual.
- Dulay, H.C. et al. (1982): *Language Two*. Oxford: Oxford University Press.
- Gilquin, G., S. De Cock & S. Granger. Forthcoming. *The Louvain International Database of Spoken English Interlanguage. Handbook and CD-ROM*. Louvain-la-Neuve: Presses Universitaires de Louvain.
- Granger, S. (2003): "Error-tagged learner corpora and CALL: a promising synergy". *CALICO Journal*, 20(3): Special Issue on Error Correction in Computer-Assisted Language Learning, 465-480.
- James, C. (1998): *Errors in Language Learning and Use. Exploring Error Analysis*. London: Longman.
- Kämmerer, S. (2009): "Error annotation from the learner's perspective: learner-centred identification and tagging of errors as a new annotation level". Paper presented at the 30th annual conference of the International Computer Archive of Modern and Medieval English (ICAME30). Lancaster, UK.
- Kämmerer, S. (forthcoming): "Interference in advanced English Interlanguage: scope, detectability and dependency", Proceedings to the bi-annual Teaching and Language Corpora Conference (TALC), Brno (Czech Republic), 30/06-03/07/2010.

# The German LINDSEI Team

---

- þ **Project leader:** Joybrato Mukherjee
- þ **Coordination:** Christiane Brand, Simone Müller
- þ **Interviews:** Rosemary Bock, Jane Dewhurst, Kristen Jung, Kathryn Khairi-Taraki, Colin Oakley
- þ **Transcriptions:** Miriam Kaemmerer, Susanne Kaemmerer, Simone Tausend, Christian Woppowa
- þ **Error tagging:** Rosemary Bock, Susanne Kaemmerer, Joybrato Mukherjee
- þ **Research:** Christiane Brand, Sandra Götz, Susanne Kaemmerer, Joybrato Mukherjee

**Contact:** Sandra.Goetz@anglistik.uni-giessen.de  
Susanne.Kaemmerer@anglistik.uni-giessen.de