Joint meeting of the EALTA SIGs
Assessment of Writing and Assessment for Academic Purposes
University of Bremen/Germany; November 18, 2017

A corpus-based approach to operationalizing and assessing writing proficiency in the academic register:

The case of reporting verbs

Marcus Callies

Roadmap

1. Learner corpora in Language Testing and Assessment

2. Describing a data-driven approach

3. Case study on reporting verbs

4. Conclusion

1. Learner corpora in LTA (1)

- learner corpus: systematic collection of **authentic, continuous and contextualized language use** by foreign/second language learners, stored in electronic format (**uncontrolled, open-ended**)
- used in SLA research for almost two decades; **Learner Corpus Research (LCR) =** computer-aided approach to storage and processing of (mostly written) learner data
- enables collection and analysis of large amounts of data (more difficult in earlier SLA research that largely used experimental data) in order to ...

  … improve in-depth description of (advanced) interlanguages

  … give SLA theories a more solid empirical foundation (alongside with experimental data)

  … help produce tools and teaching materials designed for needs of specific learner populations

  … **inform/complement assessment of L2 proficiency**

Granger, Meunier & Gilquin (2015)

1. Learner corpora in LTA (2)

- LCR responding to limitations of 'traditional' ways of assessing writing proficiency in LTA:
  - writing tasks (as part of tests)
  - expert raters
  - rating using scales
  - assigning CEFR-level

- subjectivity and variability of human rating vs. "objective", quantifiable linguistic descriptors
- recent research strand using learner corpora to inform, validate, and advance L2 proficiency assessment based on CEFR:
  - can-do statements do not provide language-specific, fine-grained linguistic details regarding learners' skills in certain registers
  - need to identify corpus-based, quantifiable linguistic descriptors ('criterial features') to add "grammatical and lexical details of English to CEFR's functional characterisation of the different levels" (Hawkins & Filipovic 2012: 5)

Taylor & Barker (2008), Hawkins & Buttery (2009, 2010), Barker (2010, 2013), Hawkins & Filipović (2012)

1. Learner corpora in LTA (3)

- three approaches:
  - corpus-informed
  - corpus-based
  - corpus-driven
- distinction based on three aspects:
  - how corpus data are actually put to use
  - aims and outcomes for LTA
  - degree of involvement of researcher in data retrieval, analysis and interpretation

→ no strict distinctions, may overlap and merge in some practises

Callies, Diéz-Bedmar & Zaytseva (2015), Callies & Götz (2015)

Corpus-informed approach

use of data	aims & outcomes for LTA	involvement of researcher	example
corpus as reference source; provides practical information on learners' language use (= what they can do) at certain levels of proficiency	evidence to inform and validate test content and practices	high	*Cambridge Learner Corpus* (CLC; Hawkey & Barker 2004); *Pearson International Corpus of Academic English* (Ackermann, Biber & Gray 2011)

Hawkey & Barker (2004), Barker (2013)

Corpus-based approach

use of data	aims & outcomes for LTA	involvement of researcher	example
corpus as source of data for linguistic research, testing existing hypotheses about learner language	evidence used to identify set of distinct features or descriptors for differentiating proficiency levels (criterial features)	medium	CLC vs. *British National Corpus* (Hawkins & Filipović 2012)

Hawkins & Buttery (2009, 2010), Hawkins & Filipović (2012)

Corpus-driven approach

use of data	aims & outcomes for LTA	involvement of researcher	example
corpus driven = data driven (Francis 1993); no preconceptions/ hypotheses prior to corpus analysis; using computer techniques for data extraction and evaluation	evidence of proficiency based on statistical analyses largely independent of human rating	low	*International Corpus of Learner English* (Wulff & Gries 2011)

Wulff & Gries (2011), Carlsen (2012)

2. Describing a data-driven approach

Aims

- identify and operationalize fine-grained linguistic descriptors for assessment of writing proficiency in academic register
- data-driven approach to assessing proficiency partially independent of human rating
- "sophisticated language use in context" = implementing and operationalizing set of "positive linguistic properties" to determine what learners can do at advanced levels when writing for academic purposes
- combining three approaches in use of learner corpora for proficiency assessment: **corpus-informed**, **corpus-based**, and **corpus-driven**

Ortega & Byrnes (2008), Callies, Diez-Bedmar & Zaytseva (2014)

Step 1

- select linguistic feature(s) that characterize academic prose (**informed by corpus research** on expert native-speaker usage)
- select descriptors in terms of **keyness**, **operationalizability** and if they remain problematic even for highly proficient L2 learners ("**late acquired features**")
- possible candidates for linguistic descriptors of academic writing:
  - specific constructions (verb-argument constructions, e.g. focus constructions, raising);
  - inanimate subjects (e.g. *This paper discusses…*, *The results suggest that…*)
  - phrases to express rhetorical functions (e.g. *by contrast*, *to conclude*, *in sum*)
  - reporting verbs (e.g. *discuss*, *claim*, *suggest*, *argue* etc.)
  - lexical co-occurrence patterns (e.g. *conduct*, *carry out*, *undertake* as typical verbal collocates of *experiment*, *analysis*, *research*)

Biber & Conrad (2009), Callies (2008, 2009, 2013), Granger & Paquot (2009), Paquot (2010)

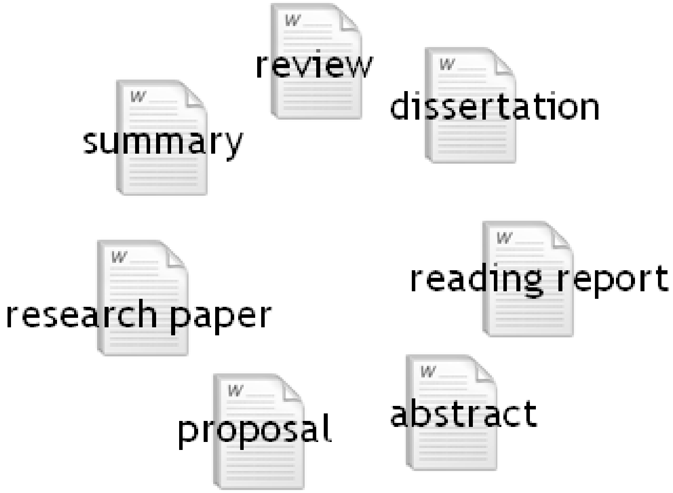
Steps 2 & 3

- retrieve and analyse descriptors in academic learner writing (**corpus-based**)
- classify and assess written sample using statistical techniques (**corpus-driven**)

The corpus

CALE
Corpus of Academic Learner English

- **Language for Specific Purposes learner corpus** containing discipline- and genre-specific texts; focus on **written academic English** ("academic learner writing")
- seven **academic text types** ("genres") produced as assignments in content courses by university students of English
- 50-100 texts per genre and L1 component
- L1 backgrounds so far: mostly German, Lithuanian, Russian, Polish, Turkish



Callies & Zaytseva (2013)

Example: 'Agentivity' of academic writing (1)

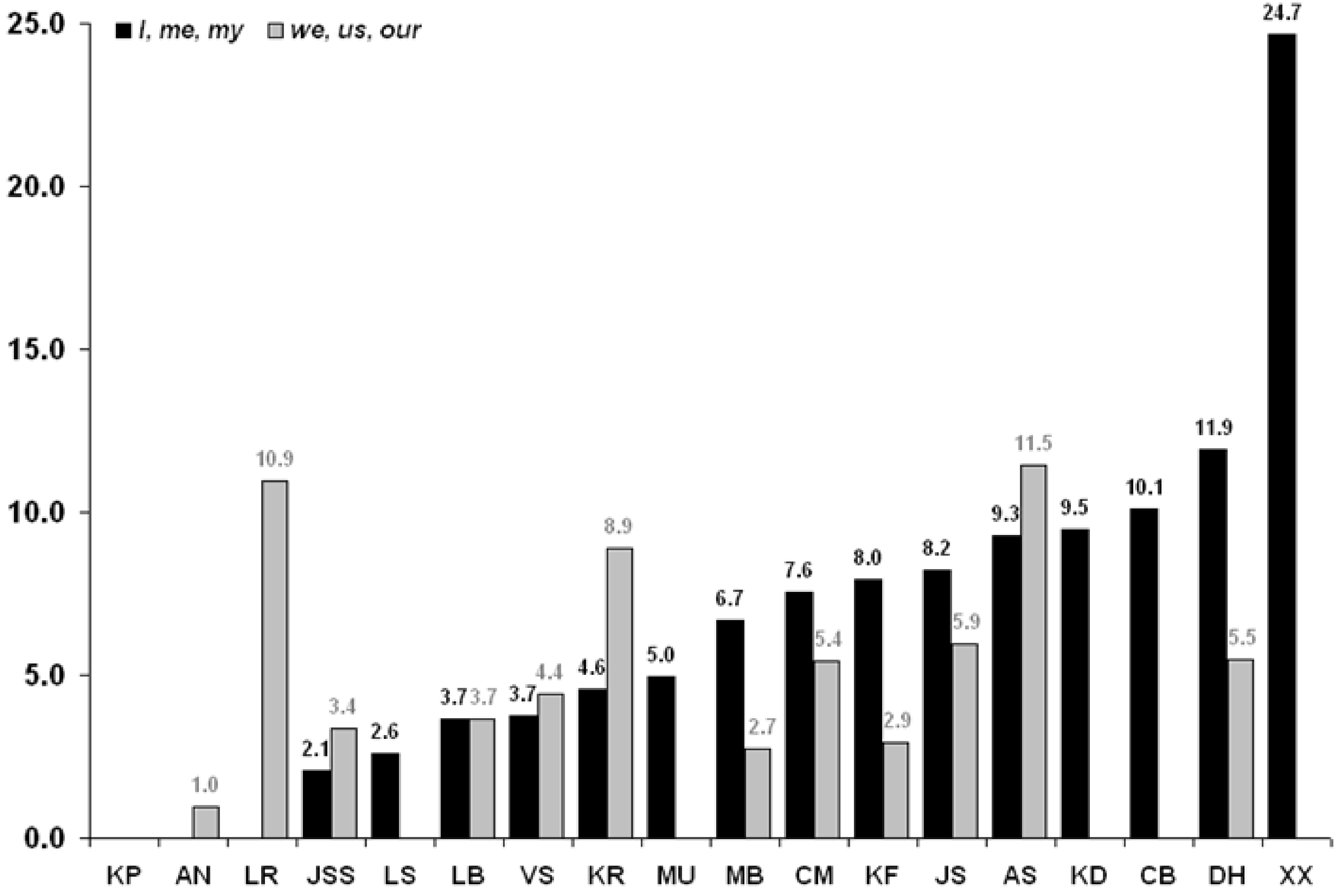
- lexico-grammatical choices to refer to writer (= author-agent) and others in **reporting events** in academic texts:

1) We will then analyze data specific to the case of Nicaraguan Sign Language to compare …
2) Neville and Buckingham (1996) and later Hyland (1999) analyze some of these linguistic options for citation in detail ...
3) To understand why this failure occurred, one must consider two factors: the experimental design and ...
4) This paper will analyze the organizational structure of art-historical discourse through a study of seven texts about portrait paintings …
5) Overall it could be argued that men have a higher social status than women due to using better forms of language.
6) Once a patient’s language abilities have been analyzed using linguistic assessment tests …

Example: 'Agentivity' of academic writing (2)

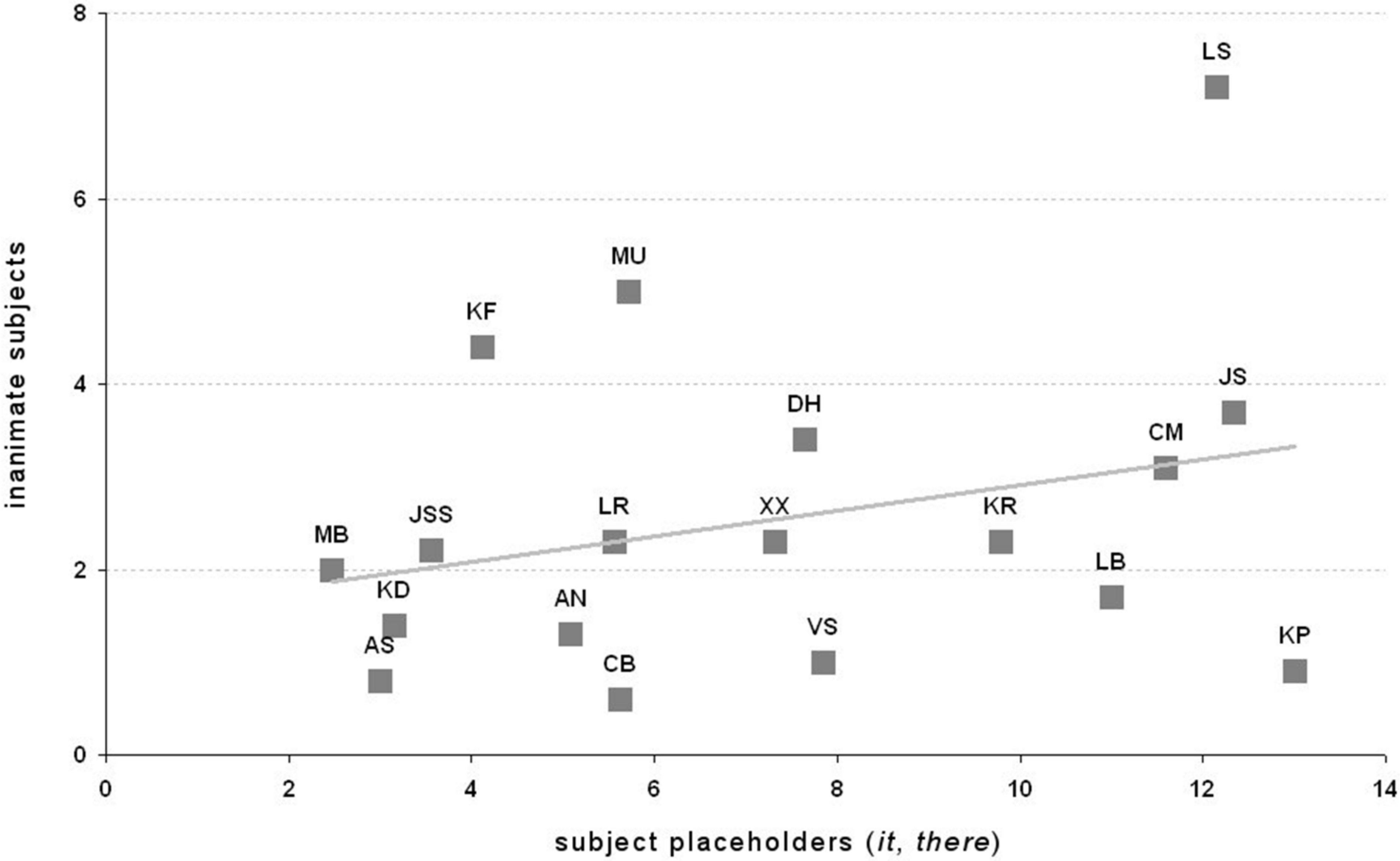
- Biber & Conrad (2009: 162): almost **no first person references in modern article**s, but **agentless passives** and **inanimate subjects** common
- CALE: 18 single-authored linguistics term papers (62,300 words, approx. 2,940 sentences) written by German EFL learners
- compared with similar subset from *Michigan Corpus of Upper-Level Student Papers* (MICUSP; Römer & O'Donnell 2011)
- main findings:
  - significant **underrepresentation of inanimate subjects in L2 writing** (but preferred reporting strategies in L1 academic English)
  - **overrepresentation of strategies to suppress agent** (due to avoidance of 1. person pronouns), e.g. **passive constructions** with semantically 'empty' subject-placeholders:
    There are two things to be discussed in this section.
    It can be observed that …

First person pronouns



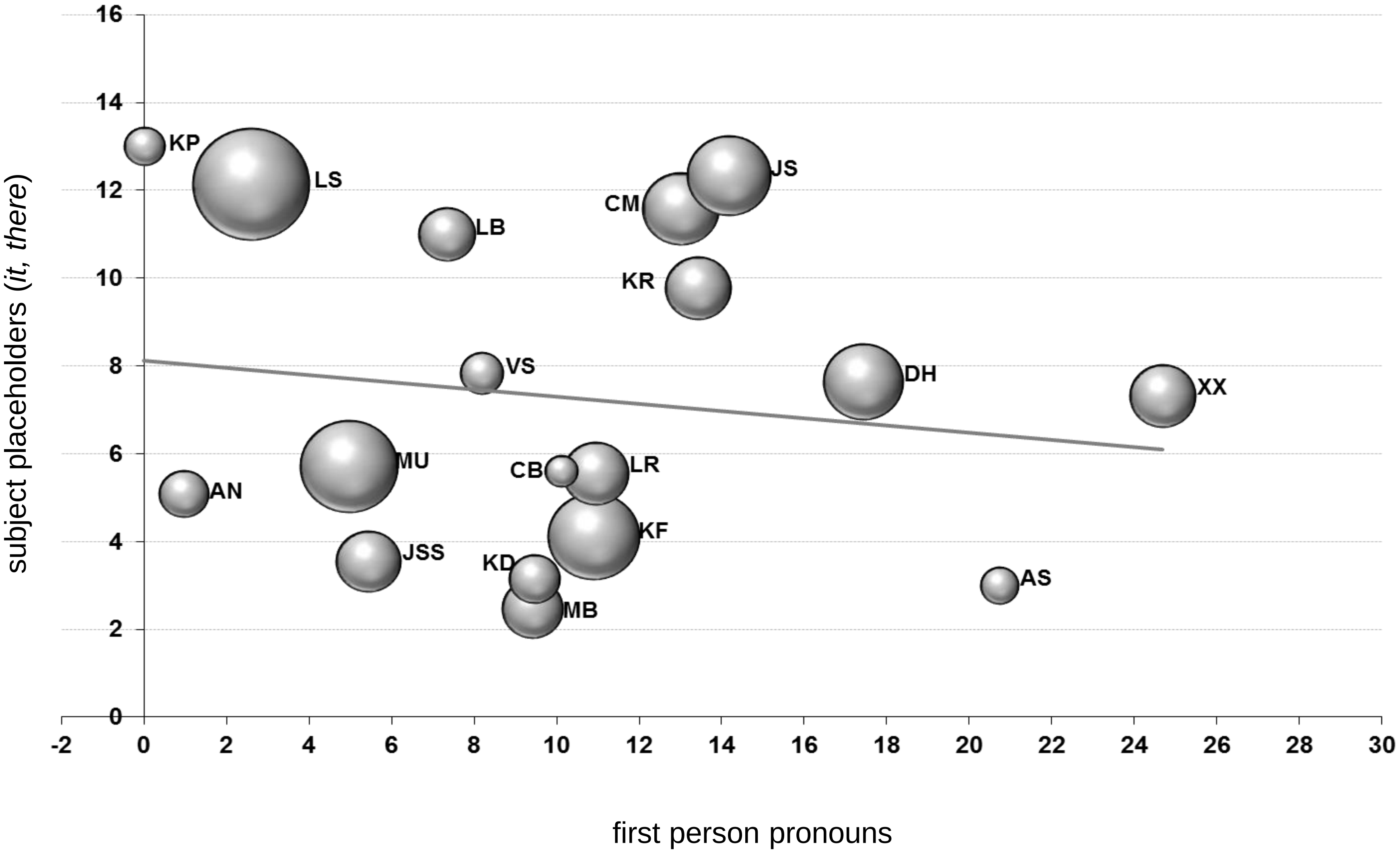
Callies (2013)

Subject placeholders and inanimate subjects



Callies (2013)

First person pronouns, subject placeholders and inanimate subjects

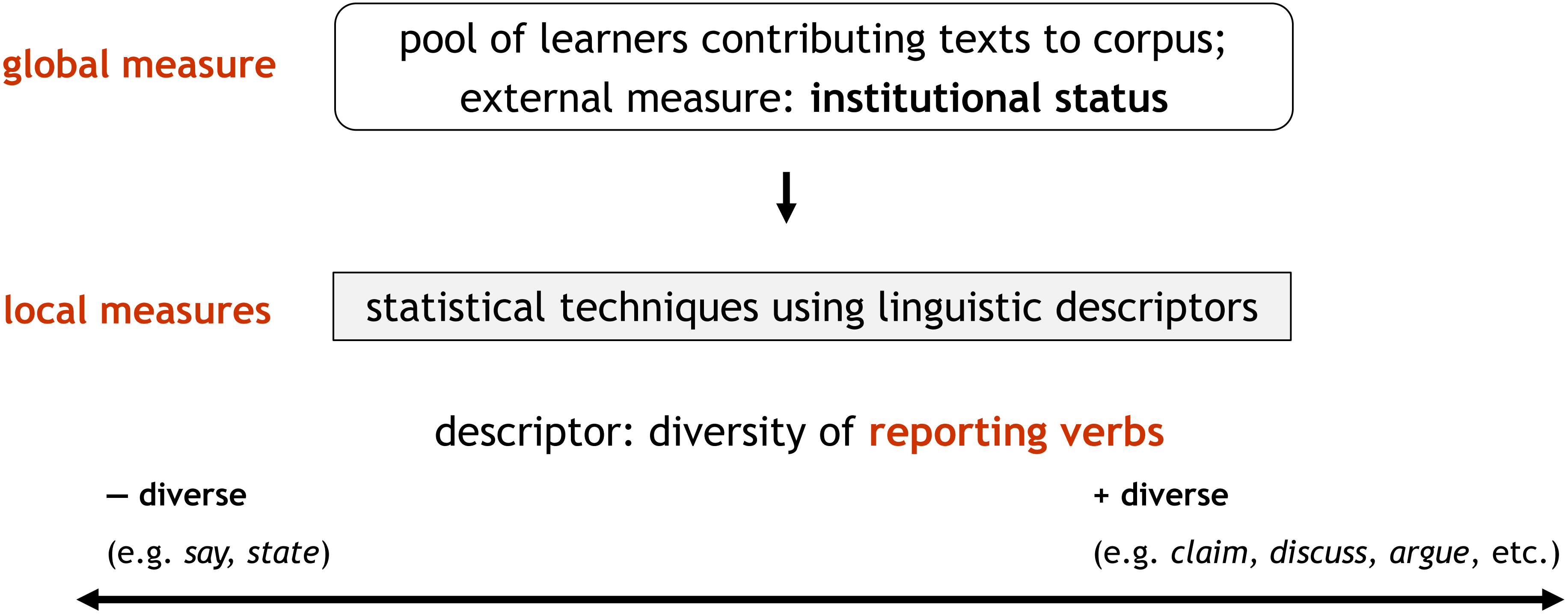


Callies (2013)

3. Case study: Reporting verbs in academic writing

- descriptor: lexical verbs frequently used to report facts and findings in academic writing = **reporting verbs** (aka research verbs, discourse verbs)
- crucial for reporting content, establishing other authors' and writer's own claims and situating these within published research
  = **high keyness**
- list of frequent reporting verbs drawn up from research literature; verbs extracted from corpus semi-automatically
  = **comparatively easy to operationalize**
- learners demonstrating high level of general language proficiency have limited inventory of reporting verbs in academic writing
  = **late-acquired**

From external criteria to linguistic descriptors



global measure

pool of learners contributing texts to corpus;
external measure: **institutional status**

local measures

statistical techniques using linguistic descriptors

descriptor: diversity of **reporting verbs**

– diverse

(e.g. *say, state*)

+ diverse

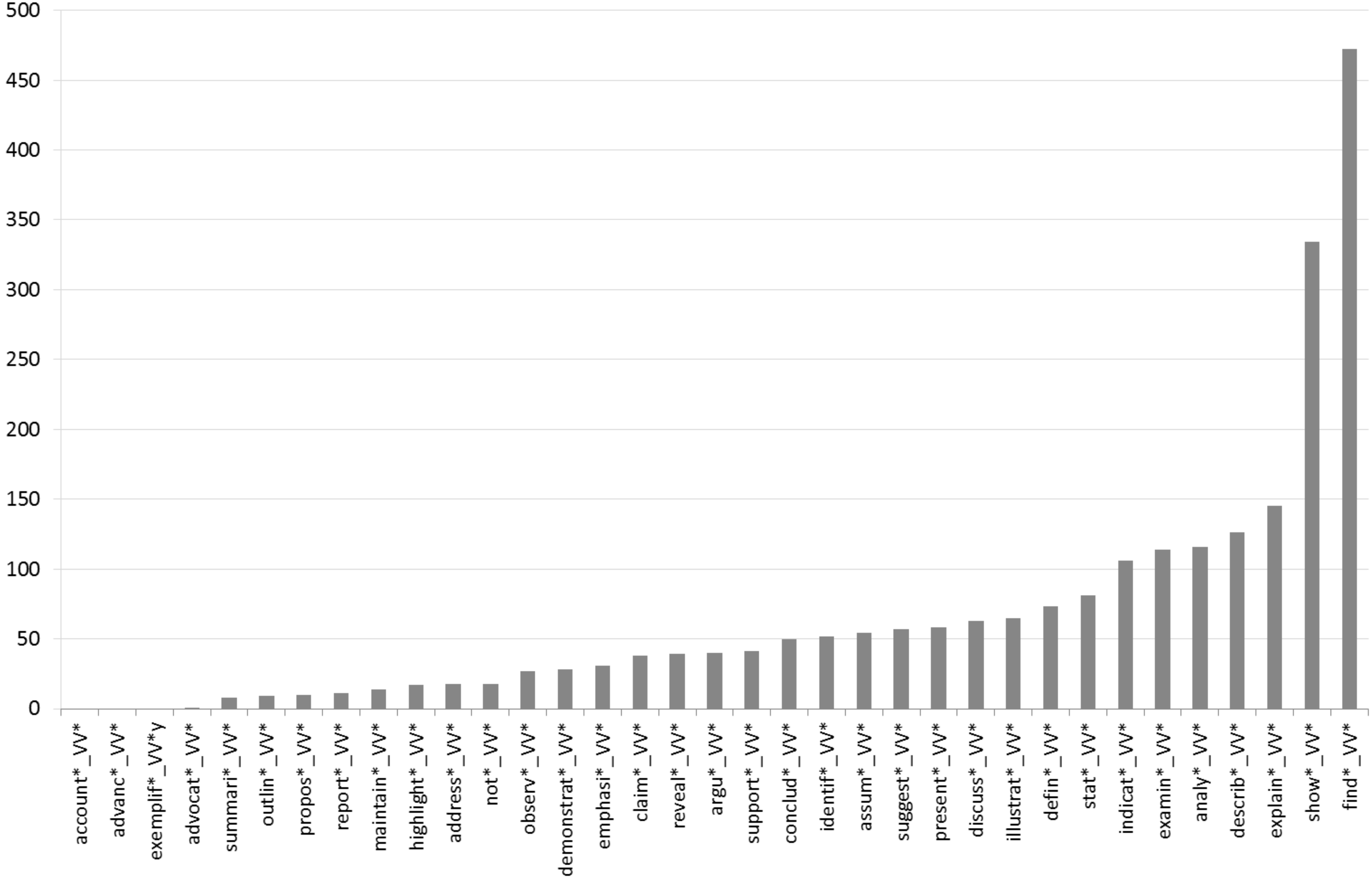
(e.g. *claim*, *discuss*, *argue*, etc.)

3. Case study: Methodology

- data: **50 research papers** produced by German EFL student writers at university; POS-tagged for corpus processing
- list of **35 frequent reporting verbs in academic writing** compiled from academic word lists and research literature
- texts fed into ***AntConc***, a corpus processing tool; target verbs extracted and counted (semi-)automatically
- texts ranked according to **diversity** of reporting verbs used:
  - size (number of tokens)
  - richness (number of types)
  - evenness (degree to which tokens are distributed equally across types)
- Simpson's Index of Diversity *D*: measure of diversity accounting for both richness and evenness; a figure between 0 and 1 = the greater the value, the greater the sample diversity

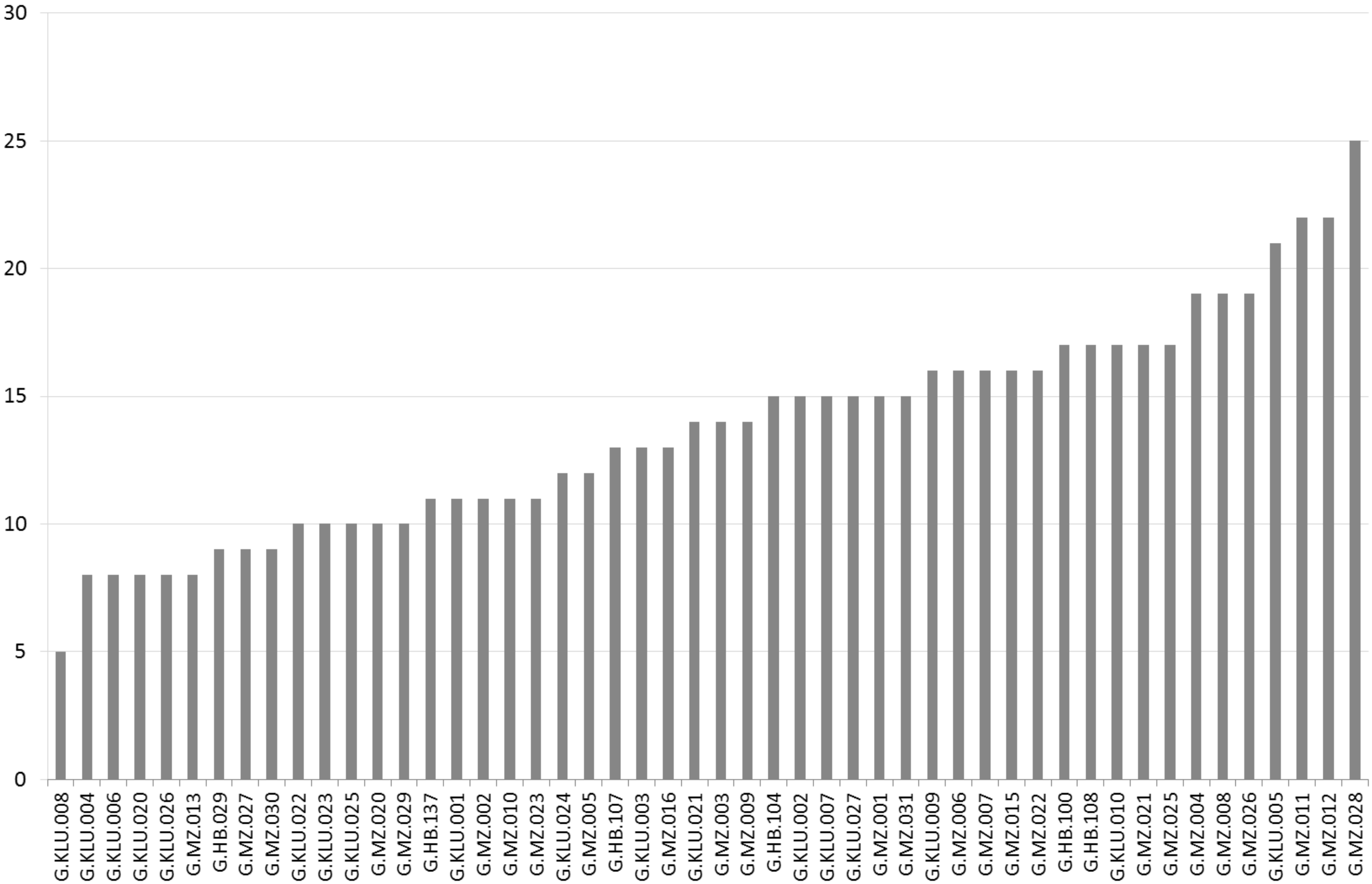
Anthony (2016), Coxhead (2000), Gardner & Davies (2014), Hyland (2002), Jarvis (2013), Paquot (2010)

3. Case study: Results (all verb tokens)



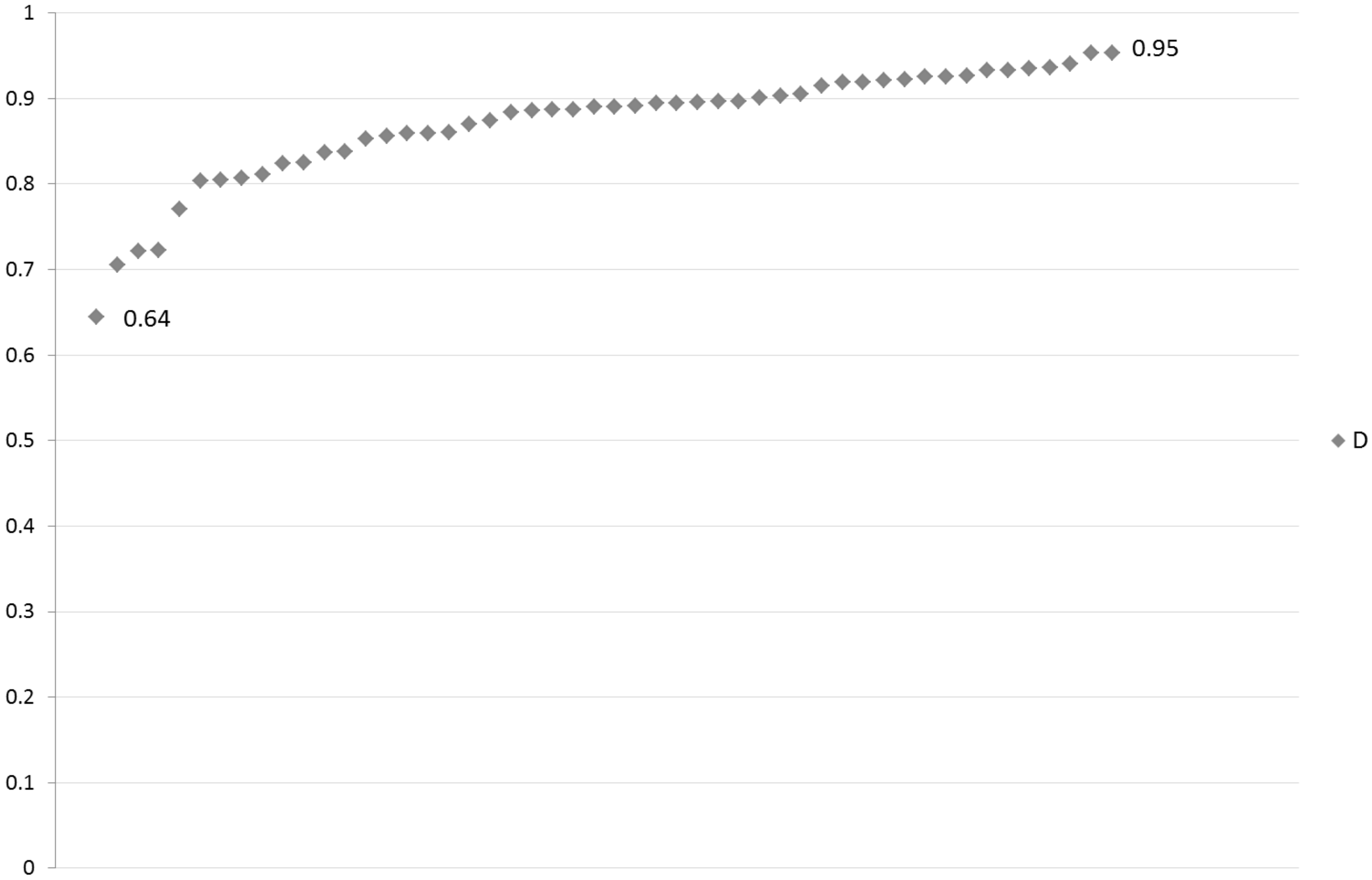
46 tokens on average / per text; SD = 22

3. Case study: Results (verb types per text)

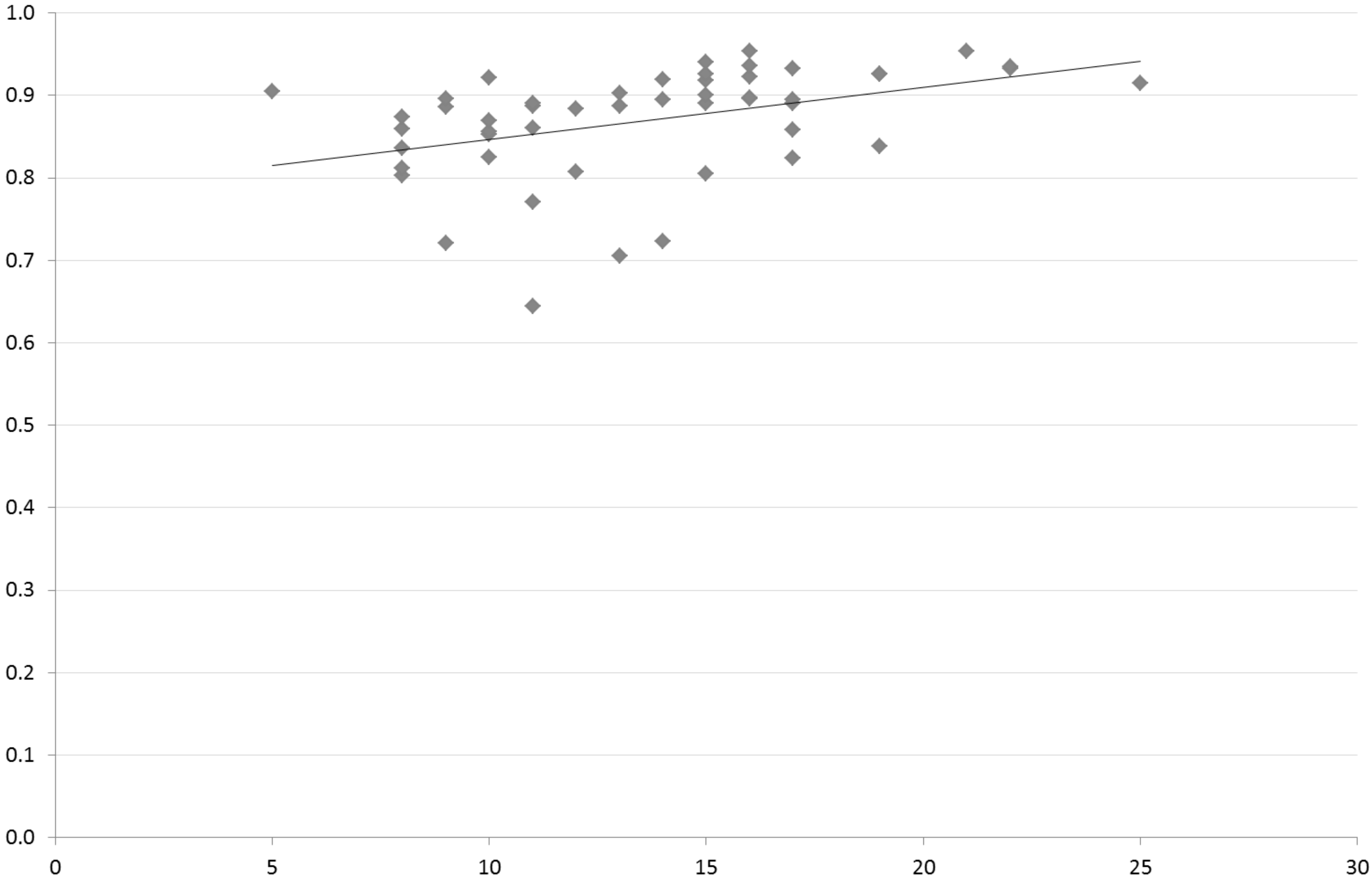


range = 20

3. Case study: Results (texts by *D*-score)



3. Case study: Results (types correlated with *D*)



Intermediate positive correlation between types and *D* (r=0.40)

4. Conclusion

- learner corpora can inform, complement and possibly advance assessment of L2 proficiency vis-á-vis the CEFR
- identify and operationalize set of descriptors that are
  - language- and register-specific
  - quantifiable
  - subject to (semi-) automatic processing
- data-driven assessment of writing proficiency in academic register involving three steps:
  1. select linguistic features for academic prose in terms of keyness, operationalizability and "late acquisition" (**corpus-informed**)
  2. retrieve descriptors from corpus (**corpus-based**)
  3. classify and assess proficiency using statistical techniques (**corpus-driven**)

References

Ackermann, K., Biber, D. & B. Gray (2011). An academic collocation list. Paper presented at *Corpus Linguistics 2011*, 20-22 July 2011, Birmingham, UK.

Anthony, L. (2016). *AntConc* (Version 3.4.4). Tokyo: Waseda University. Available from http://www.laurenceanthony.net/

Barker, F. (2010). How can corpora be used in language testing? In A. O'Keeffe & M. McCarthy (eds.), *The Routledge Handbook of Corpus Linguistics*. New York: Routledge, 633-645.

Barker, F. (2013). Assessment and testing: Overview. In C.A. Chapelle (ed.), *The Encyclopedia of Applied Lin*guistics. New York: Blackwell, 1360-1366.

Biber, D. & S. Conrad (2009). *Register, Genre, and Style*. Cambridge: Cambridge University Press.

Callies, M. (2008). Easy to understand but difficult to use? Raising constructions and information packaging in the advanced learner variety. In G. Gilquin, M.B. Diez-Bedmar & S. Papp (eds.), *Linking Contrastive and Learner Corpus Research*. Amsterdam: Rodopi, 201-226.

Callies, M. (2009). *Information Highlighting in Advanced Learner English. The Syntax-Pragmatics Interface in Second Language Acquisition*. Amsterdam: Benjamins.

Callies, M. (2013). Agentivity as a determinant of lexico-grammatical variation in L2 academic writing. *International Journal of Corpus Linguistics* 18(3), 357-390.

Callies, M., Diez-Bedmar, M. B. & E. Zaytseva (2014). Using learner corpora for testing and assessing L2 proficiency. In P. Leclercq, H. Hilton & A. Edmonds (eds.), *Proficiency Assessment Issues in SLA Research: Measures and Practices*. Clevedon: Multilingual Matters, 71-90.

Callies, M. & S. Götz (2015). Learner corpora in language testing and assessment: Prospects and challenges. In M. Callies & S. Götz (eds.), *Learner Corpora in Language Testing and Assessment*. Amsterdam: Benjamins, 1-9.

Callies, M. & E. Zaytseva (2013). The *Corpus of Academic Learner English* (CALE) – A new resource for the assessment of writing proficiency in the academic register. *Dutch Journal of Applied Linguistics* 2:1, 126-132.

Carlsen, C. (2012). Proficiency level – a fuzzy variable in computer learner corpora. *Applied Linguistics* 33:2, 161-183.

Coxhead, A. (2000) A new academic word list. *TESOL Quarterly* 34: 213-238.

Francis, G. (1993). A corpus-driven approach to grammar; principles, methods and examples. In M. Baker, G. Francis & E. Tognini Bonelli (eds.), *Text and Technology: In Honour of John Sinclair*. Amsterdam: Benjamins, 137-156.

Gardner, D. & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics* 35:3, 305-327.

Granger, S. & Paquot, M. (2009). Lexical verbs in academic discourse: A corpus-driven study of learner use. In M. Charles, D. Pecorari, & S. Hunston (eds.), *Academic writing. At the interface of corpus and discourse*. London: Continuum, 193-214.

Granger, S., F. Meunier & G. Gilquin, eds. (2015). *Cambridge Handbook of Learner Corpus Research*. Cambridge: CUP.

Hawkey, R. & F. Barker (2004). Developing a common scale for the assessment of writing. *Assessing Writing* 9, 122-159.

Hawkins, J. & P. Buttery (2009). Using learner language from corpora to profile levels of proficiency: Insights from the English Profile Programme. In L. Taylor & C.J. Weir (eds.), *Language Testing Matters: Investigating the Wider Social and Educational Impact of Assessment*. Cambridge: Cambridge University Press, 158-175.

Hawkins, J. & P. Buttery (2010). Criterial features in learner corpora: Theory and illustrations. *English Profile Journal* 1:1, 1-23.

Hawkins, J. & L. Filipović (2012). *Criterial Features in L2 English*. Cambridge: Cambridge University Press.

Hyland, K. (2002). Activity and evaluation: Reporting practices in academic writing. In J. Flowerdew (ed.), *Academic Discourse*. Harlow: Longman, 115-130.

Jarvis, S. (2013). Capturing the diversity in lexical diversity. *Language Learning* 63: Suppl. 1, 87-106.

Ortega, L. & H. Byrnes (2008). The longitudinal study of advanced L2 capacities: An introduction. In L. Ortega, & H. Byrnes (eds), *The Longitudinal Study of Advanced L2 Capacities*. New York: Routledge/Taylor & Francis, 3-20.

Paquot, M. (2010). *Academic Vocabulary in Learner Writing*. London: Continuum.

Römer, U. & M. B. O'Donnell (2011). From student hard drive to web corpus (Part 1): The design, compilation and genre classification of the *Michigan Corpus of Upper-level Student Papers* (MICUSP). *Corpora* 6:2, 159-177.

Taylor, L. & F. Barker (2008). Using corpora for language assessment. In E. Shohamy & N.H. Hornberger (eds.), *Encyclopedia of Language and Education. 2nd Edition, Volume 7: Language testing and assessment*. New York: Springer, 241-254.

Wulff, S. & Gries, S. (2011). Corpus-driven methods for assessing accuracy in learner production. In P. Robinson (ed.), *Second Language Task Complexity: Researching the Cognition Hypothesis of Language Learning and Performance*. Amsterdam: Benjamins, 61-87.