Academic text types in the Corpus of Academic Learner English (CALE)

The *Corpus of Academic Learner English* (CALE) is a learner corpus for the description of advanced learner varieties (ALVs) as to written academic English ("academic learner writing"). While existing learner corpora, such as the *International Corpus of Learner English* (ICLE), include learner writing of a general argumentative, creative or literary nature, and thus, not academic writing in a narrow sense, CALE comprises various academic text types produced in university courses of English. CALE will include texts produced by university students of English in linguistics, literary and cultural studies.

Academic writing/prose in our context is conceived of as "any writing that fulfills a purpose of education in a college or university (…); writing in response to an academic assignment, or professional writing that trained 'academics' – teachers and researchers – do for publications read and conferences attended by other academics" (Thaiss & Zawacki (2006: 4). In addition, the following features are characteristic of academic writing (Department of Translation Studies, University of Tampere/Finland):

- it represents structured research written by scholars for other scholars (with all university writers being scholars in this context)
- it addresses topic-based research questions of interest to anyone who is seeking factually-based, objectively-presented information on a particular topic
- its objective is the creation of new knowledge via (a) a review of what is currently known about a given topic as (b) the foundation for the author's new views or perspectives on the topic.

Biber et al. (1999), Biber (2006), and Biber & Conrad (2009) consider academic prose "a very general *register*, characterized as written language that has been carefully produced and edited, addressed to a large number of readers who are separated in time and space from the author, and with the primary communicative purpose of presenting information about some topic" (Biber & Conrad 2009: 32; our emphasis). Register is used as a "cover term for any variety associated with a particular configuration of situational characteristics and purposes. Thus, registers are defined in nonlinguistic terms" (Biber & Conrad 2001: 175). At the same time, they make clear that "there are usually important linguistic differences among registers as well" (Biber & Conrad 2001: 175) that correspond to the differences in situational characteristics. Biber et al. (1999) in their *Longman Grammar of Spoken and Written English*

distinguish between four major registers according to several situational characteristics (Table 1). Academic writing is characterized by a lack of interactiveness, online production, and shared immediate situation, its main communicative purpose being information, argumentation and explanation, its specialist audience and global dissemination.

	CONVERSATION	FICTION	NEWS	ACADEMIC
mode	spoken	written (+ written dialogue)	written	written
interactiveness and online production	yes	(restricted to fictional dialogue)	no	no
shared immediate situation	yes	no	no	no
main communicative purpose/content	personal communication	pleasure reading	information/ evaluation	information/ argumentation / explanation
audience	individual	wide-public	wide-public	specialist
dialect domain	local	global	regional/ national	global

Table 1. Major situational differences among four primary registers (Biber et al. 1999: 16)

Biber and Conrad (2009: chapter 5) present a more detailed description of academic prose and its situational characteristics and linguistic features. According to Biber & Conrad (2009: 32), more specialized written academic registers (sub-registers) may differ along various parameters, e.g. intended audience, communicative purpose, and academic discipline. We will refer to these as text types or genres.

The selection of text types to be included in CALE is partially modeled on two existing potential native-speaker control corpora that contain texts from L1 writers of similar academic standing (i.e. university students): the *Michigan Corpus of Upper-Level Student Papers* (MICUSP), and the *British Academic Written English* corpus (BAWE). The paper classification developed for MICUSP (Römer & Brook O'Donnell 2011) features seven paper categories: 1) argumentative essay, 2) creative writing, 3) critique/evaluation, 4) proposal, 5) report, 6) research paper, 7) response paper. BAWE includes as many as 13 "genre families", i.e. classes of genres sharing functional and structural properties (Heuboeck, Holmes & Nesi 2009: 46-50).

Some of the paper categories and genre families identified in these corpora are outside of the scope of texts types to be included in CALE, either because they represent rather specialized genres (like some included in BAWE), or they may not be considered academic in a narrow sense (like argumentative and creative essay writing in MICUSP). Our classification adopts

some of the situational characteristics and linguistic features identified for academic prose by Biber and Conrad (2009) and tries to set up clear profiles for seven categories envisaged to be represented in CALE. For that purpose, we consider the situational characteristic of communicative purpose/goal of a text a practicable criterion to set the different academic genres apart from each other. Similar to Biber and Conrad's (2009) classification we distinguish between A) a text's general purpose, B) its specific purpose(s), C) the skills the author demonstrates, and D) the author's stance.

In addition, we list the major features of each text type as to A) structural features, B) length, and C) functional features.

TEXT TYPE / GENRE	COMMUNICATIVE GOAL/PURPOSE A. General purpose / B. Specific purpose(s) / C. Skills / D. Stance	FEATURES A. Structural / B. Length / C. Functional
1 **BA/MA dissertation** **DISS**	A. informational – inform, explain, interpret B. presents detailed account of original research (or replication / extension of previous research) in view of state-of-the-art C. author demonstrates ability to carry out theoretical and/or empirical research that may include developing a research design, as well as collecting, filtering, analyzing and critically interpreting data vis-à-vis one or more research questions; author demonstrates ability to present results in an organized, meaningful way D. author's opinion/evaluation not usually overt, but may occur in literature review	A. structured into predictable sections (usually with subheadings); may include the following structural elements: abstract, introduction, literature review, methods, results, discussion, conclusion B. varies C. entire text serves to answer one or more research questions; contains original data, or compiles existing data for the purpose of providing new interpretation(s)
2 **Research paper** **RPA**	A. informational – inform, explain, interpret B. presents piece of original research (small case study or replication / extension of previous study) C. author demonstrates ability to carry out theoretical and/or empirical research that may include developing a research design, as well as collecting, filtering, analyzing and critically interpreting data vis-à-vis one or more research questions; author demonstrates ability to present results in an organized, meaningful way D. author's opinion/evaluation not usually overt, but may occur in literature review	A. structured into predictable sections (usually with subheadings); may include the following structural elements: abstract, introduction, literature review, methods, results, discussion, conclusion B. shorter than dissertation / thesis; varies C. entire text serves to answer one or more research question; contains original data, or compiles existing data for the purpose of providing new interpretation (s)
3 **Reading report** **RR**	A. informational – inform, describe, (assess) B. critically describes structure, content and significance of published research C. author demonstrates ability to understand and summarize complex text coherently D. author's opinion/evaluation may be present, but is not foregrounded and does not appear to drive text	A. not structured into sections (no subheadings) B. short, approx. 600 words C. most space is devoted to description rather than critical assessment; not driven by an original thesis or research question

TEXT TYPE / GENRE	COMMUNICATIVE GOAL/PURPOSE A. General purpose / B. Specific purpose(s) / C. Skills / D. Stance	FEATURES A. Structural / B. Length / C. Functional
4 **Summary** **SUM**	A. informational – describe B. summarizes content of published research C. author demonstrates ability to understand and summarize complex text coherently D. author's opinion/evaluation absent	A. not structured into sections (no subheadings) B. shorter than reading report; varies C. entirely descriptive, no critical assessment; not driven by an original thesis or research question
5 **Abstract** **ABS**	A. informational – inform B. captures the essence of published research (i.e. the why, how, and what, e.g. research focus, methodology results/findings, conclusion and recommendations); should help reader to quickly ascertain purpose, content and usefulness of publication C. author demonstrates ability to extract and provide essential information in an exhaustive and compelling way D. author's opinion/evaluation absent	A. not structured into sections; appears at beginning of text it accompanies; may also occur as stand-alone entity instead of full paper B. rather short (approx. 100-250 words), rarely exceeding 500 words C. self-contained piece of writing, can be understood independently from accompanying publication
6 **Review** **REV**	A. informational – inform, describe and evaluate/assess B. presents brief descriptive summary and evaluation/assessment of effectiveness, validity, or usefulness of published research; may offer recommendations for improvement C. author demonstrates ability to understand significance of publication and to evaluate/assess its quality D. author's opinion/evaluation foregrounded, drives text	A. usually not structured into sections, but may distinguish between descriptive summary of content and evaluation (with accompanying subheadings) B. varies C. driven by an evaluation of published research as to its methodology, quality of data, findings and line of argumentation (often interwoven with descriptive account)
7 **Proposal** **PRO**	A. informational – inform, describe, argue B. proposes potential study: puts forth one or more research questions that author wishes to explore in order to further understanding of given topic; provides information about how proposed study will be tackled methodologically C. author demonstrates ability to convincingly argue for relevance, significance and manageability of proposed study D. author's opinion/evaluation not usually overt, but may occur in literature review	A. may be structured into sections with subheadings; may include the following structural elements: introduction, literature review, research questions, proposed methodology B. varies C. does not present or synthesize new data, but may include projected results; links back to relevant literature and/or previous studies; justifies the need for and outlines methods of data collection

References

Biber, Douglas (2006), *University Language. A corpus-based study of spoken and written registers*. Amsterdam: Benjamins.

Biber, Douglas & Susan Conrad (2001), "Register variation: A corpus approach", in Schiffrin, Deborah, Tannen, Deborah & Heidi Hamilton (eds.), *The Handbook of Discourse Analysis*. Oxford: Blackwell, 175-196.

Biber, Douglas & Susan Conrad (2009), *Register, Genre, and Style*. Cambridge: CUP.

Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan & Edward Finegan (1999). *Longman Grammar of Spoken and Written English*. Harlow: Longman.

Römer, Utw, & Brook O'Donnell, Matthew (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.

Heuboeck, Alois, Jasper Holmes & Hilary Nesi (2009), *The BAWE Corpus Manual version II*. http://www.reading.ac.uk/AcaDepts/ll/app_ling/internal/bawe/BAWE.documentation.pdf (accessed 04 December 2013).

Thaiss, Christopher J. & Therese M. Zawacki (2006), *Engaged Writers, Dynamic Disciplines: Research on the Academic Writing Life*. Portsmouth/NH: Boynton/Cook.

The Online Writing Lab at Purdue and Purdue University, *Common Writing Assignments*, http://owl.english.purdue.edu/owl/section/1/3/ (accessed 12 November 2010).

Department of Translation Studies, University of Tampere/Finland, *What Is 'Academic Writing'?* http://www.uta.fi/FAST/FIN/RESEARCH/acadwrit.html (accessed 12 November 2010).