Tracking Written Learner Language: Background, context and studies of the TRAWL corpus

Ingrid Kristine Hasund, Eli-Marie D. Drange, Henrik Martin Hansen Torjusen

University of Agder

1. Introduction

This special issue introduces the TRAWL (Tracking Written Learner Language) corpus and presents some of the first studies that have been carried out on the TRAWL material thus far. TRAWL is a longitudinal corpus of young learner writing in primary and secondary schools in Norway, with authentic school texts in L1 Norwegian, L2 English and the most commonly taught L3s in Norwegian schools, French, German and Spanish[1]. The aim of the TRAWL research group has been to create a corpus that addresses some of the needs in current learner corpus research, is compiled according to established design criteria and is openly available for research. With this special issue, we wish to inspire a wide range of studies of young learner writing based on TRAWL data in the years to come.

Before giving an overview of the 12 articles in this special issue (Section 3), we will provide a brief description of the background of learner corpus research (Section 1.1), followed by a presentation of some related learner corpora in international as well as Norwegian contexts (Sections 1.2-1.4). Section 2 focuses on the establishment of the TRAWL research group and on how we were able to create a longitudinal and multilingual corpus of authentic young learner writing. By sharing our hands-on experiences from the process of collecting and processing such data, we hope to encourage other researchers who would like to venture into the risky but rewarding field of creating this type of corpus. Section 3 presents the content of the special issue, and finally, Section 4 is dedicated to acknowledgements.

---

[1] For a definition of how the terms L1, L2 and L3 are used in this special issue in connection to the TRAWL corpus, see Section 1.3.

1.1 Background: Corpus linguistics and learner corpus research

Corpus linguistics, the study of language based on computerised databases of written or spoken texts, was established as an area of research in the 1960s (see e.g. Biber et al., 1998; Sinclair, 1991). Learner corpus research (henceforth, LCR), however, did not emerge as a field until the end of the 1980s, when academics and publishers started acknowledging the potential of using learner corpora for research and teaching purposes (Callies & Paquot, 2015, p. 1). Learner corpora can be analysed using advanced linguistic software tools which make it possible to study many different aspects of learner language. Over the years, learner corpus data and corpus linguistic methods have gradually been applied in Second Language Acquisition (SLA) research as well as Foreign Language Teaching (FLT) research, although the development towards a closer collaboration between LCR on the one hand and SLA/FLT research on the other has been rather slow (Granger, 2002, p. 6; 2021, p. 8; Myles, 2015, pp. 48–49).

Learner corpora are compiled according to strict design criteria and vary along several different dimensions (for an overview, see Gilquin, 2015), including target language (L2/L3) and L1 backgrounds, time and scope of collection, medium, and genre. Depending on the design criteria used in the collection process, each learner corpus is annotated with information (metadata) about learners and texts. Metadata about learners often include age, gender, language background and (expected) proficiency level, while text metadata regularly include medium, genre, topic, task type and task conditions. In addition, learner corpora are provided with linguistic annotation, such as error tagging and word-class, or part-of-speech (POS) tagging, making it possible to carry out detailed searches for particular linguistic features and constructions (see e.g. van Rooy, 2015).

The first learner corpora, such as the International Corpus of Learner English (ICLE)[2], focused primarily on written L2 English, reflecting the dominant position of written English in corpus linguistics generally. Later, other languages "progressively joined the learner corpus bandwagon" and more spoken and multimedia corpora were also being compiled (Granger, 2012, p. 12). This development has continued to the present, as manifested in the list of learner corpora around the world published by the Centre for English Corpus Linguistics, CECL (2022), where one finds an increasing number of learner corpora from languages other than English (including multilingual corpora) as well as spoken and multimedia corpora. Regarding time of collection, most learner corpora are cross-sectional, with learner texts gathered at a single point in time. Longitudinal corpora, tracking the same learners over a period of time,

---

[2] https://uclouvain.be/en/research-institutes/ilc/cecl/icle.html

have long been called for (see e.g. Jarvis & Pavlenko, 2008, p. 230; Meunier, 2015, pp. 381–382) but are still comparatively rare. Of the 198 learner corpora listed by CECL (2022), only 17 are described as longitudinal. When it comes to scope of collection, the majority of available corpora have focused on advanced learners from the university level, with few corpora consisting of young learner writing. Finally, regarding genre, the majority of existing learner corpora are restricted to one or a few genres, with the argumentative essay clearly dominating. In recent years, however, more corpora have been compiled that include a wider variety of genres (Gilquin, 2015, p. 12; CECL, 2022).

1.2 Corpora of young learner writing worldwide

As shown above, there is a dearth of available young learner corpora, and the TRAWL project was initiated to address this need. Before detailing the Norwegian context in which TRAWL was created, we will briefly describe what other written corpora exist worldwide that have a design similar to TRAWL. The International Corpus of Crosslinguistic Interlanguage, ICCI[3] (cf. Tono & Díez-Bedmar, 2014) is a cross-sectional corpus of EFL texts from eight different countries/regions (Hong Kong, Germany, Israel, China, Japan, Poland, Singapore, Spain, and Taiwan). According to the project website (14 November 2022, cf. footnote 3), ICCI was collected 2008-2010 and comprises 6,700 argumentative and descriptive essays (around 530,000 words) written by primary and secondary school pupils with many different L1 backgrounds. Most texts are from school years 6 to 12; there are also some texts from school years 3-5.

Another corpus is SWIKO (the Swiss Learner Corpus)[4], which was collected between 2017 and 2022 in Switzerland with spoken and written data from French and German-speaking Swiss pupils in school years 10-12 (ages 14-16, mainly lower secondary level). Unlike ICCI, which focuses on L2 English only, SWIKO is a multilingual corpus consisting of texts in L1 German and French and in L2 German, French and English. The written part of the corpus consists of around 1,900 texts and 116,000 words (Karges et al., 2022; T. Studer, personal communication, 27 October 2022; Studer & Hicks, 2022). Both SWIKO and ICCI comprise cross-sectional data, with texts written in response to specifically designed tasks in eliciting argumentative and narrative/descriptive genres.

---

[3] http://corpus-hong.shisu.edu.cn/icci/index.jsp
[4] https://centre-plurilinguisme.ch/en/research/swiss-learner-corpus-swiko

Finally, LEONIDE (the Longitudinal lEarner cOrpus iN Italiano, Deutsch, English)[5] is a multilingual corpus from the Italian province of South Tyrol, where Italian and German are the official languages. Collected between 2015-2018, the corpus consists of 2,512 texts (approximately 240,000 words) written by lower secondary school pupils (ages 11-14) in their L2 German and Italian and L3 English, plus reference data written by the same pupils in their L1 Italian or German. LEONIDE resembles SWIKO and ICCI in that it used genre-specific prompts created by the researchers. It differs, however, from SWIKO and ICCI in that it is longitudinal; the pupils were asked to write similar texts (a picture story and an opinion text) in each of the three consecutive years of lower secondary school (Glaznieks et al., 2022).

1.3 The Norwegian context

This section takes stock of written corpora from Norway that are available for research. In the description of the Norwegian school system in this special issue, school years 8-10 are used for lower secondary school, and since most pupils will turn 13 when entering the first year of lower secondary school and 15 when entering the last year, the age group for this level is defined to be 13-15[6]. For the upper secondary level, school years 11-13 are used, and the age range is 16-18. For primary school, the school years are 1-7, and the ages are 6-12.

Since we work with languages used in a Norwegian school context, we refer to Norwegian as the L1 because it is the first language introduced in school year 1 and the main language of instruction; for most pupils it is also their first language. The second language introduced in school (also in year 1) is English, which we define as the L2; French, German and Spanish are all introduced later (in year 8) and will be referred to as the L3. There are, of course, pupils who know other languages in addition to these and for which these languages are not strictly speaking the first, second and third language to be learnt, but in this special issue, we shall use the terms L1 Norwegian, L2 English and L3 French, German and Spanish in overall descriptions of the TRAWL data.

Although our main focus is on languages other than Norwegian, we will begin by presenting existing corpora of Norwegian for two reasons: firstly, because Norway has a fairly long tradition of compiling corpora of written Norwegian from younger pupils and several corpus-based

---

[5] https://www.porta.eurac.edu/lci/leonide/

[6] This definition follows established practice in Norway and is in line with that used by The Norwegian Directorate for Education and Training (udir.no). Evidently, some pupils have not yet turned 13 when they enter lower secondary school in year 8 and some will have turned 16 by the time they leave year 10, so the actual age range is 12-16.

studies have been published over the years (see below); secondly, because such corpora and studies are relevant from a contrastive perspective. TRAWL also includes some L1 Norwegian texts for comparison with the L2 English and L3 French, German and Spanish material.

To the best of our knowledge, the oldest available corpus of written Norwegian from a school context is the KAL (Kvalitetssikring av Læringsutbyttet i Norsk Skriftlig "Quality Assurance of the Learning Outcome in Written Norwegian") corpus from lower secondary school[7], collected 1998-2001. The KAL project aimed to assess the learning outcome in written Norwegian at the end of lower secondary school (Berge et al., 2003; Berge, 2005) and part of the project was to compile a corpus of authentic pupil texts from the year 10 school-leaving exam. Texts were collected over four consecutive years (each year from a new group of pupils) and the KAL corpus contains around 3,300 exam scripts with grades.

Another relatively old corpus is ASK[8] (Norsk Andrespråkskorpus "The Norwegian Second Language Corpus"), which was compiled in the early 2000s. It contains authentic texts written by L2 learners of Norwegian with several different L1s in a language test for adult immigrants (Tenfjord, Meurer & Hofland, 2006). ASK has approximately 770,000 words in 1,936 texts, and has been used to study the influence of the learners' L1s on their acquisition of L2 Norwegian (see e.g. Golden et al., 2017).

A more recent corpus is SKRIV (Skriving i Videregående skole "Writing in Upper Secondary School")[9], collected in 2012. The SKRIV project aimed at studying and developing pupils' writing and text competence in the vocational study programmes at upper secondary school (Selj & Ryen, 2020). The SKRIV corpus consists of authentic pupil texts from various subjects, amounting to around 225 texts and 112,000 words. The project focused specifically on pupils with Norwegian as an L2 or L3, but the corpus also contains some texts from pupils with Norwegian as their L1.

Finally, the research project NORM (Developing National Standards for the Assessment of Writing – a Tool for Teaching and Learning)[10], which was carried out 2012-2016, studied writing and assessment practices across subjects in Norwegian primary school (Berge et al., 2019). As part of the project, a corpus of authentic pupil texts and teacher assessments was collected 2012-2014. Texts were collected from school years 3, 4, 6 and 7 and the NORM

---

[7] https://www.hf.uio.no/iln/om/organisasjon/tekstlab/prosjekter/kal/index.html

[8] https://clarino.uib.no/corpuscle

[9] https://www.hf.uio.no/iln/om/organisasjon/tekstlab/prosjekter/skriv/

[10] https://www.hf.uio.no/iln/om/organisasjon/tekstlab/prosjekter/norm/index.html

corpus consists of 5,200 written texts amounting to around 1.1 million words. While the other Norwegian corpora mentioned in this section mainly contain cross-sectional data, NORM has both cross-sectional and longitudinal data in that some of the pupils were followed over two years.

With regard to written corpora from languages other than Norwegian, there are some L2 English corpora that have been created in Norway. Although based on older learners from the university level, two corpora are worth mentioning in this context because of their link to international learner corpus research: ICLE-NO and VESPA-NO are the Norwegian components of the international corpora ICLE (see footnote 2) and VESPA (Varieties of English for Specific Purposes dAtabase)[11]. Both are written corpora with English L2 texts produced by Norwegian learners and have been used in a number of studies of Norwegian advanced learner English (See e.g. Hasselgård & Johansson, 2011; Paquot et al., 2013; Hasselgård, 2015; Larsson et al., 2020; Granger & Larsson, 2021; Ebeling & Hasselgård, 2021; Hasund & Hasselgård, 2022). ICLE-NO consists of 317 argumentative essays (about 214,000 words) collected between 1999 and 2002 (Johansson, 2008, p. 116; Granger et al., 2020, p. 33). VESPA-NO was collected between 2009 and 2018 and comprises 388 academic essays (approximately 530,000 words) from different disciplines[12]. As TRAWL has been compiled following many of the same design criteria as ICLE-NO and VESPA-NO, these corpora are suitable for contrastive studies of younger versus older learners. It should be mentioned, though, that both corpora are cross-sectional and based on argumentative/academic essays written in response to prompts provided by the researchers. Regarding learner corpora based on L3 languages in Norway, there are, as far as we know, only one available corpus, namely the ELENOR (Español Lengua Extranjera en NORuega) corpus of written texts in L3 Spanish from university students[13].

When it comes to young learner writing, there is only one available corpus in Norway besides TRAWL, the CORYL (Corpus of young learner language) corpus of L2 English, collected 2004-2005. It has 706 texts (around 340,000 words) written by Norwegian pupils in school years 7, 10 and 11 in the course of the National Testing of English[14]. The texts are error-coded and are also assigned levels on the CEFR scale, which makes the corpus especially suitable for research on writing development (cf. Hasselgreen & Sundet, 2017). CORYL resembles ICLE-

---

[11] https://uclouvain.be/en/research-institutes/ilc/cecl/vespa.html

[12] https://www.hf.uio.no/ilos/english/services/knowledge-resources/vespa/index.html

[13] https://www.hf.uio.no/iln/tjenester/kunnskap/sprak/korpus/skriftsprakskorpus/elenor/index.html

[14] https://clarino.uib.no/corpuscle

NO and VESPA-NO in that it is cross-sectional, but differs in that the texts are authentic responses to tasks given in the National Testing of English and comprise a number of different text types/genres.

1.4 Summing up

As this introduction has demonstrated, there is currently no accessible corpus of young learner writing that is truly longitudinal, is based on authentic school writing in a variety of genres and tracks the same learners in up to three languages (L1, L2 and L3); TRAWL is thus the first corpus to follow this particular design. The next section is devoted to a presentation of how the TRAWL team has worked to create a longitudinal and multilingual corpus of authentic young learner writing. The technical details of the corpus are described in Dirdal et al. (2022); here, we will primarily focus on telling the 'story of TRAWL'. One of the main reasons why there are so few corpora like TRAWL is that they are very difficult to compile, requiring a huge amount of time, planning, resources and endurance. As Meunier (2015, p. 381) expresses it, "[such] difficulties in collecting data mean that the high demand for longitudinal learner corpora is – quite unsurprisingly – met with few research teams collecting such data types". By sharing our story of how the project came about and how we have worked to carry it through, we hope to inspire and encourage other research teams to embark on similar projects in the future.

2. The story of TRAWL

The TRAWL project was initiated by Associate Professor Hildegunn Dirdal at the University of Oslo's (UiO) Department of Literature, Area Studies and European Languages, in 2013. Attending the 34th ICAME conference[15] in Spain that year, Dirdal was inspired by presentations about learner corpora and discussed the idea with colleagues from Inland Norway University of Applied Sciences (INN) who were also present at the conference. Together with colleagues at UiO and the Norwegian University of Science and Technology (NTNU), the group decided to embark on a project compiling a longitudinal learner corpus in Norway. Dirdal had her background in the field of Second Language Acquisition, and the other group members all worked in teacher training, had done research on learner language, and had close links to numerous partner schools around Norway. The group secured funding for a small pilot, compiling

---

[15] https://www.usc.es/en/congresos/icame34/

English and French learner data, and started work on an application for a larger project, for which Professor Anne-Line Graedler at INN was chosen as PI.

At about the same time, a similar initiative had been launched independently by a group of researchers at the University of Agder's (UiA) Department of Foreign Languages and Translation. The corpus initiative was named Elevspråk i Transitt (ESIT), which means 'pupils' language in transit', as the group was interested in tracking the writing development of young L2/L3 learners in the transition from lower secondary to upper secondary school. The ESIT project included staff members from UiA who were involved in teacher training in L2 English and L3 Spanish and German. It was first led by Professor Signe Mari Wiland, whose background included research on L2 English teaching at various levels, and later by Associate Professor Eli-Marie D. Drange, who had a long experience with corpus compilation and corpus linguistic research on L1 Norwegian and Spanish through her work on the Norwegian part of the UNO corpus[16] and the Spanish COLA corpus[17].

In 2015, Graedler and her colleagues approached the ESIT team at UiA with an invitation to develop a joint application for large-scale external funding from the Research Council of Norway. The application was submitted the same year, and, while awaiting the result, the group applied for and was granted internal, small-scale funding from all four universities involved, to recruit student assistants and start data collection at schools in different parts of Norway. Dirdal was head of the fieldwork at UiO. At NTNU, the fieldwork was led by Associate Professor Tale Margrethe Guldal from the Department of Teacher Education, who had a solid background in English teacher training and had done research on pupils' written L2 English. At INN, the fieldwork was led by Professor Anne-Line Graedler at the Faculty of Education. Graedler had many years of experience in learner corpus compilation and L2 English research through her participation in the NEST[18] corpus project, where she was project leader, and the Norwegian component of the LINDSEI[19] corpus.

Our first application for large-scale funding to the Research Council of Norway was rejected; with internal funding, however, we were already up and running and in 2016, TRAWL was

---

[16] UNO: Språkkontakt og Ungdomsspråk i Norden – Talespråkskorpuset, Oslodelen 'Language Contact and Teenage Language in Scandinavia – The Norwegian Spoken Corpus of Oslo Teenage Language'.

[17] COLA: Corpus de Lenguaje Adolescente 'Corpus of [Spoken] Teenage Language'.

[18] NEST: Norwegian English Student Translations, see http://korpus.uib.no/humfak/nest/index-e.html

[19] LINDSEI: The Louvain International Database of Spoken English Interlanguage, see https://uclouvain.be/en/research-institutes/ilc/cecl/lindsei.html

officially established as a national research group in a collaboration between UiO, INN, NTNU and UiA. From 2016 to 2019, the group was led by Graedler (INN), who is credited with giving the corpus its name, "Tracking Written Learner Language", as well as the acronym TRAWL; from 2019 to the present, it has been led by Dirdal (UiO). At UiA, ESIT changed its name to TRAWL-UiA in 2016 and also changed from being a project to being a research group for two reasons. One was that the ESIT team had grown to become the largest of the TRAWL divisions and thus needed on-site management at UiA. Another reason was that its new status as a separate and locally administered research group triggered administrative funding and increased opportunities to seek internal financing. Since 2016, TRAWL-UiA has been led by Professor Ingrid Kristine Hasund at the Department of Foreign Languages and Translation. Hasund, who was also part of the former ESIT team, worked in L2 English teacher training and had experience from corpus compilation and research through her work on the COLT corpus[20], which is now part of the British National Corpus (BNC),[21] and the Norwegian part of the UNO corpus (cf. footnote 16).

A few more words about funding are required. The collection and processing of longitudinal learner data is, to say the least, "a highly time-consuming and resource-intensive task", as Myles (2015, p. 330) expresses it. External research institutions such as the Research Council of Norway do not provide funding for data collection and processing alone, only as part of larger research projects – which are indeed hard to get, as we will return to below. For the most part, we have had to rely on internal funding in the form of small and medium-sized grants from the departments/faculties at our respective universities. Such grants do not go a long way, but it is rare for universities to commit to the stable, long-term internal funding required to complete the collection of longitudinal data. As stated by Meunier (2015, p. 381), "research funders do not like to commit resources for very long periods of time". And surely, we have been met with a concerned "When will you finish?" many times as we have kept pushing back the end date for our project. But this is how we were able to continue employing and training a constantly changing pool of student assistants over so many years: To relentlessly apply for small and medium-sized grants from any internal funding source that was available, to persistently plead for support ("Just one more year now!") and to never, ever give up. That is to say: As a team, TRAWL has never given up. Individual team members have come and gone for a number of

---

[20] COLT: the Bergen Corpus of London Teenage Language, see
https://www.nb.no/sprakbanken/en/resource-catalogue/oai-clarino-uib-no-colt/
[21] http://www.natcorp.ox.ac.uk/

reasons, but most team members have remained since our initial proposal for large funding from the Research Council in 2015. As mentioned, this application was rejected, as were all subsequent applications we submitted in the years that followed. It was not until 2021 that the TRAWL team finally succeeded in securing large-scale financing: The research project *MULTIWRITE – Interactions Between First, Second and Third Languages*, which is led by Dirdal and involves a collaboration between UiO, UiA, the University of Bergen and the University of Exeter in England, received four years of funding to investigate pupils' writing development, feedback from the teachers and teacher collaboration across language subjects[22].

To round off this section, we would like to give some reflections on what we find has been most crucial to TRAWL's success, namely the team itself and the relationships with our partners. As for the team itself, one important success factor is that the team members together possess a broad and solid background from SLA, FLT and corpus linguistic research, and have hands-on experience with corpus compilation from other projects. Another success factor, which is also a potential source of friction, has been accepting that each team member has their own research interests and agendas. Our project spans several age groups and languages, and we have had to work continually to find and maintain unity as a research team, working towards a common goal while at the same time nurturing individual projects and ideas. Regarding our partners, the fact that most team members are involved in teacher training means we have long-standing and solid relationships with a number of local praxis partner schools, which has made it easier for us to continually recruit pupils and teachers to provide us with data (see Dirdal et al., 2022, for details about the data material), as well as to keep recruiting and training teacher students as student assistants. Without the diligent work of our student assistants in collecting, anonymizing, coding and organizing the data, the corpus could never have been finalized. For several years, we have had weekly workshops for assistants and staff members, first on campus and later online. This way of including students in a research group entails training them for research, which is part of our responsibility as academic staff. As the assistants have learned the necessary skills, they have passed on their knowledge by training new assistants entering the project. A further aspect of students' research training is learning how to write and publish scientific work, and several student assistants have written their MA theses based on TRAWL data.

---

[22] https://www.hf.uio.no/ilos/english/research/projects/multiwrite/index.html

3. Contents of the special issue

The special issue consists of a main section and a forum section. The main section contains eleven double-blind peer-reviewed articles, while the forum section contains one article written by a former MA student. The presentation of the articles in the main section is divided into four parts. The first part contains an introductory article co-authored by Dirdal, Hasund, Drange, Vold and Berg, which presents the design and construction of the TRAWL corpus. The second part contains three articles based on L3 data; Drange writes about Spanish, Hamann about German and Vold about French learner texts. The third part has four articles on L2 English data, written by Berg, Hasund, Nacey and Rørvik. The fourth part comprises three articles by Dirdal, Evang and Hasselgård that adopt a cross-linguistic or contrastive perspective, comparing L2 English texts with L1 Norwegian texts from TRAWL and/or L1 English texts from other corpora.

In our first article in part two, "The development of sentence complexity in the writing of young learners of L3 Spanish in Norway", Eli-Marie D. Drange explores the development of sentence complexity in the writing of young learners of Spanish as the third language (L3) learned in Norwegian schools. The study investigates texts written by learners of Spanish from the first and second year of upper secondary school (school years 11-12, ages 16-17) as their ordinary schoolwork. The findings show that there may be a discrepancy between the pupils' actual knowledge and the expected knowledge according to the reference levels. These findings support the need to develop more nuanced measures for young and novice language learners.

Also focusing on writing within language learning, Veronika Hamann's article "Writing in German as a foreign language in Norwegian upper secondary school: An investigation of patterns of language choices for meaning-making" seeks to identify and describe characteristic patterns of language choices in texts written by Norwegian upper secondary school students of German as a foreign language (age 17, school year 12, 5th year of FL learning). By studying a set of 12 learner responses to a writing prompt, Hamann aims to examine how the learners use ideational meaning-making resources to arrive at meaningful content. Overall, the study provides insights into relevant patterns for expository writing in general and points to the sophistication of the learners' language use and the linguistic demands regarding the task at hand. In line with existing research, the study also shows how SFL and genre theory can be successfully applied to the analysis of responses by beginner to intermediate GFL learners.

Eva Thue Vold's article "Development of lexical richness among beginning learners of French as a foreign language" explores the extent to which 14 upper-secondary students of

French as a foreign language (FFL) in Norwegian schools showed signs of increased lexical richness in their written production over a period of approximately six months. The study revealed that although many learners showed signs of using a somewhat more varied vocabulary over time, it is difficult to find proof of productive vocabulary development among beginning learners of FFL within such a limited period. Vold's research suggests that finer-grained measures of analysis could be added to existing automated tools to make these tools more useful for beginner levels.

The first article in part three, focusing on L2 English using the TRAWL corpus, is "Written corrective feedback in the lower secondary EFL classroom: exploring questions of what, how and why in observed and self-reported teacher practice" by Elin Maria Berg. In her article, she presents a Norwegian case study of two English teachers' written corrective feedback (WCF) provided to three students during three years of lower secondary EFL instruction. The qualitative and longitudinal approach provided insight into the complex nature of authentic learner language, implications for using such data in research, and factors affecting teachers' WCF decision-making. However, the data suggests that students do not engage with feedback enough to benefit from its learning potential.

Ingrid Kristine Hasund's article "Genres in young learner L2 English writing: A genre typology for the TRAWL (Tracking Written Learner Language) corpus" presents a genre typology for annotating learner texts from the lower secondary level in Norway (ages 13-15, school years 8-10) with data drawn from TRAWL. As no detailed genre typology exists for classifying learner texts at the lower secondary level, Hasund examines a genre typology developed by Ørevik (2019) for the upper secondary level. Hasund's analysis showed that Ørevik's (2019) genre typology was largely suitable for annotating the selected TRAWL data and only had to be slightly modified. Together with Ørevik's (2019) study, Hasund's study paves the way for more genre-based studies of L2 English writing across the lower and upper secondary levels. Furthermore, by highlighting some of the theoretical and methodological challenges with the genre typology, the analysis may inform discussions about genre in L2 English teaching.

In her article, "Development of metaphorical production in learner language: A longitudinal perspective", Susan Nacey studies the progress of five secondary school pupils aged 13-17 in Norway and the development of their metaphorical production as L2 learners of English. Central questions in Nacey's study are to assess how metaphorical density varies over time, examine the distribution of metaphor clusters over time, and explore the functions the identified metaphor clusters serve in the written discourse of these language learners. In this sense, the

aim of the longitudinal corpus-based exploration is to shed light on how metaphorical production changes as pupils progress through different semesters and grades in their school careers.

The final article on TRAWL and English L2 writing is Sylvi Rørvik's "Noun-phrase complexity in the texts of intermediate-level Norwegian EFL writers: stasis or development?". In her analysis, Rørvik examines the longitudinal development in noun-phrase complexity in English texts written by Norwegian learners in school years 8-10. With a focus on nine pupils, Rørvik tests if longitudinal development in noun-phrase complexity can be traced in the written production of intermediate-level Norwegian EFL writers in school years 8-10. Rørvik finds little evidence to suggest an increase in sophistication as regards phrasal modification over the three years. Instead, the pupils primarily rely on the types of modifiers that are acquired in early developmental stages, such as attributive adjectives and prepositional phrases.

The first article in part four is Hildegunn Dirdal's article "Cross-linguistic influence in the acquisition of relative clauses by Norwegian learners of English". In the text, Dirdal examines cross-linguistic influence in the acquisition of relative clauses by young Norwegian learners of English by comparing L1 Norwegian and L2 English material from the TRAWL Corpus to L1 English material from the Growth in Grammar Corpus. With this comparison, Dirdal investigates whether similarities between relative clause systems may lead to more subtle effects in the choice of relativizer, the type of head nominal, the syntactic function of the relativized item, the extent of relativization from embedded clauses and the use of relative clauses in special constructions such as existentials and clefts. Although the material is limited, the study found traces of the Norwegian system in the learners' L2 English, signalling that this is an area worth further investigation.

Kaja H.S.Ø. Evang's article "A Bigram-Based Exploration of Phraseological Development in Norwegian Secondary School Students' Writing in English L2" investigates the phraseological development of Norwegian intermediate L2 learners of English (aged 13 to 17). The aim is to examine association measure as a way of calculating the collocational strength and certainty of words in word pairs (bigrams) in the L2 learners' writing. Evang suggests that the pattern of association measures found for intermediate Norwegian learners of English does not match the previous findings for advanced learners. Instead, an initial decline is uncovered, in that the students produce a higher proportion of bigrams with high MI scores and high t-scores in the first year than they do one and two years later. At higher levels of proficiency, the scores increase again, the pattern resembling previous findings.

In her comparative study, "Adverb-adjective combinations in young writers' English (EL1 and EL2)", Hilde Hasselgård looks at adverb-adjective combinations in narrative writing by lower secondary school pupils in Norway and the UK. By investigating the TRAWL and Growth in Grammar Corpus corpora, Hasselgård finds a number of differences between the two writer groups. While the construction was more frequent and widespread in EL2 (English as a second language), it showed more variability in EL1 (English as a first-language) regarding syntax, semantics and lexical choice. In particular, the amplifying function of modifiers was more dominant in EL2 writing at the cost of other modifier functions. There was also a stronger concentration on a few highly frequent intensifiers in the EL2 than in the EL1 material.

In our forum section, Jovana Dasic analyses the relationship between language learning and gaming in her article "Acquiring English Through Virtual Worlds". Dasic studies the connection between Norwegian lower secondary pupils' gaming habits and their essay grades, their lexical richness in L2 English and their attitudes towards English as a Second Language (L2) language learning. Studying 14 Norwegian lower secondary pupils, with a total of 20 essays (6 from school year 8, 14 from school year 9), the results revealed a statistically significant positive correlation between the amount of time the participants spent gaming and their English essay grades. The findings also suggest that large amounts of time spent gaming are beneficial to other aspects of the student's English proficiency, such as greater self-confidence when speaking English and creativity when writing.

4. Acknowledgements

List of former and current TRAWL members and student assistants:

Staff members, including PhD and postdoc researchers:

University of Oslo

Hildegunn Dirdal

Eva Thue Vold

Kaja Evang

Mari Wærner Wesenberg

Stine Hulleberg Johansen

Anne-Kathrine Woldsnes

Inland Norway University of Applied Sciences

Anne-Line Graedler

Susan Nacey

Sylvi Rørvik

University of Agder

Ingrid Kristine Hasund

Eli-Marie Danbolt Drange

Elin Maria Berg

Veronika Hamann

Lenka Garshol

Susan Erdmann

May Olaug Horverak

Sigrunn Askland

Signe Mari Wiland

Norwegian University of Science and Technology

Tale Margrethe Guldal

Student assistants (all institutions), in alphabetical order:

Adrian Verner Krogstad

Alexandra Haver

Andre Odland

Anna Giskes

Anna Marie Kristine Hallheim Reiersøl

Anne-Kathrine Woldsnes

Cathrine Thorsland

Else Marie Gjerdsjø Risdal

Erica V. C. Brændeland

Hanne Kittelsen

Hege Henriksen

Helene Øya

Håkon Stensvand

Ida Marie Wullum

Inga Kaleck

Ingrid Karoline Dønnestad

Jens Ytre-Arne

Jovana Dasic

Julia Kristina Rebekka Vejby

Kamilla Kjøstvedt

Karola Sophie Hösterey

Kristin Opsahl

Linnea Høegh-Omdal

Magda Monteagudo

Marcus René Innvær Andersen

Marie Iversdatter Røsok

Marikken Auensen

Student assistants (cont.)

Maritza Milena Muñoz Delgado

Markus Jørgensborg

Marte Andersen

Milena Andu

Ole Hasund Mustad

Rebecca Anastacia Torkildsen

Renate Gjetnes

Renate Teigland

Robin Eriksen

Rut Hildegunn Drangsland Gutierrez

Sarah Andersen

Sigrid Graedler Listuen

Sigrid Gran

Sigurd Eye

Silje Kathrin Grønbeck

Synne A. G. Kolsvik

Torleif Kveset

Valjeta Krasniqi

Valon Sadiku

Veronica Dahlby Tveitan

Ylva Dagsdottir

References

Berge, K. L. (2005). *Ungdommers skrivekompetanse: B. 2: Norskeksamen som tekst: Vol. B. 2*. Universitetsforlaget.

Berge, K. L., Evensen, L. S., Hertzberg, F., & Vagle, W. (2003). *Ungdommers skrivekompetanse: B. 1: Norsksensuren som kvalitetsvurdering: Vol. B. 1*. Universitetsforlaget.

Berge, K. L., Skar, G. B., Matre, S., Solheim, R., Evensen, L. S., Otnes, H., & Thygesen, R. (2019). Introducing teachers to new semiotic tools for writing instruction and writing assessment: Consequences for students' writing proficiency. *Assessment in Education: Principles, Policy & Practice*, *26*(1), 6–25. https://doi.org/10.1080/0969594X.2017.1330251

Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.

Callies, M., & Paquot, M. (2015). Learner Corpus Research: An interdisciplinary field on the move. *International Journal of Learner Corpus Research*, *1*(1), 1–6. https://doi.org/10.1075/ijlcr.1.1.00edi

Centre for English Corpus Linguistics (2022): Learner Corpora around the World. Université catholique de Louvain. https://uclovain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html

Dirdal, H., Hasund, I. K., Drange, E-M., Vold, E., & Berg, E-M. (2022). Design and construction of the Tracking Written Learner Language (TRAWL) corpus: A longitudinal and multilingual young learner corpus. *Nordic Journal of Language Teaching and Learning*, *10*(2), 115–135.

Ebeling, S. O., & Hasselgård, H. (2021). The Functions of N-Grams in Bilingual and Learner Corpora: An Integrated Contrastive Approach. In S. Granger (Eds.), *Perspectives on the L2 Phrasicon: The View from Learner Corpora* (pp. 25–48). Multilingual Matters.

Gilquin, G. (2015). From design to collection of learner corpora. In S. Granger, G. Gilquin & F. Meunier, *The Cambridge handbook of learner corpus research* (pp. 9–34). Cambridge University Press.

Glaznieks, A., Frey, J.-C., Stopfner, M., Zanasi, L., & Nicolas, L. (2022). Leonide: A longitudinal trilingual corpus of young learners of Italian, German and English. *International Journal of Learner Corpus Research*, *8*(1), 97–120. https://doi.org/10.1075/ijlcr.21004.gla

Golden, A., Jarvis, S., & Tenfjord, K. (2017). *Crosslinguistic influence and distinctive patterns of language learning: Findings and insights from a learner corpus*. Multilingual Matters.

Granger, S. (2002). A Bird's-eye view of learner corpus research. In S. Granger, J. Hung & S. Petch-Tyson (Eds.), *Language Learning & Language Teaching* (Vol. 6, pp. 3–33). John Benjamins Publishing Company. https://benjamins.com/catalog/lllt.6.04gra

Granger, S. (2012). How to use foreign and second language learner corpora. In A. Mackey & S. M. Gass (Eds.), *Research Methods in Second Language Acquisition: A Practical Guide* (pp. 7–29). Wiley-Blackwell.

Granger, S. (2021). Have Learner Corpus Research and Second Language Acquisition finally met? In B. Le Bruyn & M. Paquot (Eds.), *Learner Corpus Research Meets Second Language Acquisition* (pp. 243–257). Cambridge University Press.

Granger, S., & Larsson, T. (2021). Is core vocabulary a friend or foe of academic writing? Single-word vs multi-word uses of thing. *Journal of English for Academic Purposes*, *52*, 100999. https://doi.org/10.1016/j.jeap.2021.100999

Granger, S., Dupont, M., Meunier, F., Naets, H., Paquot, M. (2020). *The International Corpus of Learner English. Version 3*. Presses universitaires de Louvain. https://hdl.handle.net/2078.1/229877

Hasselgreen, A., & Sundet, K. T. (2017). Introducing the CORYL Corpus: What it Is and How We Can Use it to Shed Light on Learner Language. *Bergen Language and Linguistics Studies*, *7*. https://doi.org/10.15845/bells.v7i0.1107

Hasselgård, H. (2015). Lexicogrammatical features of adverbs in advanced learner English. *ITL International Journal of Applied Linguistics*, *166*(1), 163-189.

Hasselgård, H., & Johansson, S. (2011). Learner corpora and contrastive interlanguage analysis. In F. Meunier, S. De Cock & G. Gilquin (Eds.), *A Taste for corpora: In honour of Sylviane Granger* (pp. 33–62). John Benjamins Publishing Company.

Hasund, I. K., & Hasselgård, H. (2022). Writer/reader visibility in young learner writing: A study of the TRAWL corpus of lower secondary school texts. *Journal of Writing Research, 3*(3), 417–472. https://doi.org/10.17239/jowr-2022.13.03.04

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. Routledge.

Johansson, S. (2008). Contrastive analysis and learner language: A corpus-based approach [Unpublished research paper]. University of Oslo. https://www.hf.uio.no/ilos/english/research/groups/English_Language_and_Corpus_Linguistics_Research/papers/

Karges, K., Studer, T., & Hicks, N. S. (2022). Lernersprache, Aufgabe und Modalität: Beobachtungen zu Texten aus dem Schweizer Lernerkorpus SWIKO. *Zeitschrift für germanistische Linguistik*, *50*(1), 104–130. https://doi.org/10.1515/zgl-2022-2050

Larsson, T., Callies, M., Hasselgård, H., Laso, N. J., Van Vuuren, S., Verdaguer, I., & Paquot, M. (2020). Adverb placement in EFL academic writing. Going beyond syntactic transfer. *International Journal of Corpus Linguistics*, *25*(2), 155–184. https://doi.org/10.1075/ijcl.19131.lar

Meunier, F. (2015). Developmental patterns in learner corpora. In S. Granger, G. Gilquin & F. Meunier (Eds.), *The Cambridge Handbook of Learner Corpus Research* (pp. 379–400). Cambridge University Press. https://doi.org/10.1017/CBO9781139649414.017

Myles, F. (2015). Second language acquisition theory and learner corpus research. In S. Granger, G. Gilquin & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* (pp. 307–331). Cambridge University Press.

Paquot, M., Hasselgård, H., & Oksefjell, S. E. (2013). Writer/reader visibility in learner writing across genres. A comparison of the French and Norwegian components of the ICLE and VESPA learner corpora. In S. Granger, G. Gilquin & F. Meunier (Eds.), *Twenty Years of Learner Corpus Research: Looking back, Moving ahead* (pp. 377–387). Presses Universitaires de Louvain.

Selj, E., & Ryen, E. (2020). *Med språklige minoriteter i klassen: Språklige og faglige utfordringer*. Cappelen Damm.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.

Studer, T. & Hicks, N. S. (2022). *SWIKO: task, lexical features and ratings* [Conference presentation]. The 31st Conference of the European Second Language Association (EuroSLA 31), University of Fribourg, Switzerland. https://eurosla31.dryfta.com/index.php

Tenfjord, K., Meurer, P., & Hofland, K. (2006). *The ASK Corpus – A Language Learner Corpus of Norwegian as a Second Language*. Proceedings from 5th International Conference on Language Resources and Evaluation (LREC), Genova 2006. https://www.lrec-conf.org/proceedings/lrec2006/pdf/573.pdf

Tono, Y., & Díez-Bedmar, M. B. (2014). Focus on learner writing at the beginning and intermediate stages: The ICCI corpus. *International Journal of Corpus Linguistics*, *19*(2), 163–177. https://doi.org/10.1075/ijcl.19.2.01ton

van Rooy, B. (2015). Annotating learner corpora. In S. Granger, G. Gilquin & F. Meunier, *The Cambridge handbook of learner corpus research* (pp. 79–105). Cambridge University Press.

Ørevik, S. (2019). *Mapping the text culture of the subject of English: Genres and text types in national exams and published learning materials*. PhD thesis, University of Bergen. http://bora.uib.no/handle/1956/19266.