

Written corrective feedback in the lower secondary EFL classroom: exploring questions of what, how and why in observed and self-reported teacher practice

Elin Maria Berg

University of Agder

elin.m.berg@uia.no

Abstract

For decades, scholarly debates have been concerned with the effect of corrective feedback (CF), both written and oral, on L2 language development. Much of the research that supports written corrective feedback (WCF) comes from short-term focused feedback studies, representing a type of feedback practice not necessarily applicable in classroom contexts. This has pointed to a need for more classroom research of authentic WCF and its effect on written learner language. Attempting to explore authentic classroom data longitudinally, this article presents a Norwegian case study of two English teachers' WCF provided to three students during three years of lower secondary EFL instruction. The student texts are part of the TRAWL (Tracking Written Learner Language) corpus and were collected from obligatory mock exams. The texts came with teacher WCF as well as revised versions. Semi-structured interviews with the teachers provided information about the teachers' practices and beliefs related to WCF. The collected data was analyzed qualitatively and later discussed against cognitive, motivational, and sociocultural theories. The process of tracking changes in error patterns confirmed some of the issues with using accuracy and global error scores as a measurement for improvement and development, interpreted as proof of learning in language acquisition research. The data further suggests that students do not engage with feedback enough to benefit from its learning potential.

Keywords

Written corrective feedback, comprehensive feedback, longitudinal classroom research, qualitative learner corpus research, EFL.

1. Written corrective feedback in the lower secondary EFL classroom: exploring questions of what, how and why in observed and self-reported teacher practice

The effect of feedback in written L2 learner development has long attracted attention in fields related to language teaching, i.e., second language acquisition (SLA), L2 writing research and composition studies. SLA studies of singular language items report clear advantages for students who receive written corrective feedback (WCF) over control groups that receive no WCF (Bitchener & Knoch, 2008, 2009, 2010; Ellis et al., 2006; Frear & Chiu, 2015), and scholars now largely agree that there is proof of positive, at least short-term, effects of WCF. Disagreements concern what the feedback should focus on, the overall amount, and which type of feedback one should use for different feedback situations. Questions like these are important since corrections may be perceived as supportive but also discouraging, leading to both positive and negative effects on the learner. Because teachers use WCF to help learners develop various aspects of their writing, not just language form, and because they rarely focus on one or two issues only, ignoring everything else, results from focused feedback research may not be as relevant from a classroom perspective (Lee, 2020). Thus, scholars have pointed to a need for longitudinal studies of authentic comprehensive classroom data, investigating the possible long-term effects of WCF (Lee, 2020; Storch, 2010; Van Beuningen et al., 2012). Aiming to address this need, this case study explores longitudinal authentic teacher WCF provided to authentic written learner language. The setting is a lower secondary EFL classroom in Norway, and the study follows three students and two teachers during three years of instruction. A complete presentation of the material and analysis can be found in Berg (2020).

The following research questions were formed:

- RQ1. To what extent and how do EFL teachers provide WCF on lexical, grammatical, mechanical, and sentence structure errors to the same students' written texts in grade 8, 9, and 10?
- RQ2. Tracking different types of spelling and verb tense errors, does the students' accuracy improve in subsequent writing following WCF they receive and revise?
- RQ3. What beliefs about WCF and errors guide the two teachers' practice?

To answer these questions, this article will first introduce some key terms and definitions, present the central ideas of the theories used in the discussion, and review literature relevant to the field of teacher WCF practice. The methods section will then describe the design of the study as well as decisions made for categorizations and analysis. Results from the error and feedback

analysis concerning RQ1 and RQ3 are presented alongside the interview data in two parts: one for each teacher. This is followed by a presentation of the subsequent analyses made to answer RQ2. Lastly, the discussion will tie the various parts together, leading to concluding thoughts, implications, and suggestions for future research.

1.1 Terms and definitions

Written corrective feedback (WCF) is a teacher's written response to student writing (Mao & Crosthwaite, 2019). It provides information about aspects a student can change to make a text better or more accurate. Teachers may make direct changes in the text, write suggestions in the margin, use different color codes, circle/underline, or write comments at the end with explanations and suggestions for improvement. WCF may target errors and issues *directly* (providing both location and a corrected version or suggestion), *indirectly* (pointing out but not providing a solution), or be provided with added metalinguistic explanations (Ellis, 2008b). An error analysis provides further information about the type of error targeted, whether local (word-level) or global (sentence- or text-level), easily correctable, or more cognitively challenging. The number of error categories affects the level of detail in the analysis and may be adapted to describe linguistic categories as well as teacher feedback focus (Eckstein & Ferris, 2017; Ellis & Barkhuizen, 2005; Ferris, 2006). To describe what a learner has done wrong or left out to produce the error, one may also apply a surface structure analysis (Dulay, Burt, & Krashen, 1982). Literature sometimes refers to WCF as corrective feedback (CF) or written CF. In this article, the three variations are used interchangeably.

Teacher WCF tends to focus on grammatical errors but is sometimes used for lexical and non-grammatical errors (e.g., punctuation or spelling) (Bitchener & Storch, 2016). If the teacher or researcher focuses on a few selected types of errors, generally between one and five, s/he uses a *selective* or *focused* approach. In contrast, corrective feedback to many or all errors is termed *comprehensive* or *unfocused* (Lee, 2020)¹. Lastly, the amount of feedback is often calculated in feedback points, where each point is pointing the learner's attention to an error/issue that may be corrected or changed.

Because WCF specifies what students can improve, or errors to correct, it is necessary to define what an error is. One definition is that errors in L2 production are "deviations from the norms of the target language" (Ellis, 2008a, p. 51). Another definition is that errors and mistakes

¹ For an extensive discussion about when to use which term and why this may be challenging, see Lee (2020).

miscommunicate the writer's intended meaning (Bitchener & Ferris, 2012, p. 42). Both are subject to interpretation, and it is up to the teacher to evaluate the students' individual need for feedback. Researchers study errors in learner language, searching for signs of learning or feedback effects. Measurements of accuracy, such as global error scores (the number of errors divided by the total number of words), are often used to examine patterns of change. Decreasing numbers of errors indicate improved accuracy, which is interpreted as proof of learning. When using such methods to measure language development, considerations about where to draw the line between errors and mistakes become important, as this will have an impact on the number of errors counted. This, however, is not always straightforward, seeing that learners tend to simultaneously make mistakes and errors, sometimes due to being tired, unfocused, or sloppy, sometimes because of partial competence, and sometimes because they lack the competence (Corder, 1967; Ellis, 2008a). To avoid misinterpretation of ambiguous cases, this study considers all errors and mistakes as errors only.

2. Theory

2.1 Cognitive theories

Cognitive theories in SLA, such as the Skill Acquisition theory and the Information Processing Model, are based on the notion that learning is a cognitive processing of information or input (e.g., Anderson, 1983; McLaughlin, 1987, 1990). The consolidation process of learning is defined as the processing that must occur when learners attempt to understand or apply feedback points in revision or new pieces of writing (Bitchener, 2019).

Krashen (1984) maintains that the level of conscious attention affects the learning process. He made a distinction between learned and acquired competence. While acquired competence is the result of exposure to natural language, learned competence is attained through deliberate attention to target language (TL) rules. According to Ellis (2008b), these two competencies are built on distinct types of linguistic knowledge: implicit (for acquisition) and explicit (for learning). Theorists disagree as to whether explicit knowledge can become implicit knowledge. Those in favor believe it is possible through abundant practice and feedback. Bitchener & Storch (2016) suggests that the nature of written CF, with its unlimited time for processing, indicates that "CF has the potential to facilitate the development of explicit L2 knowledge and, through practice, the acquisition of implicit knowledge" (p. 12).

The "noticing hypothesis" (Schmidt, 1995, 2001) and the "output hypothesis" (Swain, 1995) are also relevant for discussions of WCF. In these theories, WCF is facilitative of learning

because it stimulates noticing and noticing-the-gap (Sheen, 2010, p. 170). Learners engage in output hypothesizing as they communicate and this interaction together with CF eventually leads to acquisition.

2.2 Sociocultural theory

In Sociocultural theory, or SCT (Vygotsky, 1978, 1981), learning is seen as a social phenomenon where the social factor is constitutive of cognition rather than just an important variable in learning processes (Villamil & Guerrero, 2006). SCT suggests that everything we learn appears in social interaction before it becomes learner internalized. Development can be traced as the learner moves from being dependent on the expert to being an independent user of abstract concepts. This takes place through language and appropriate forms of scaffolding (Bitchener & Storch, 2016, pp. 68–69). WCF can function as a successful type of scaffolding when it targets a learner's Zone of Proximal Development (ZPD), tailored to the learner's individual stages of development (Sheen, 2010, p. 170). This aspect of SCT suggests that individual differentiation is necessary when choosing feedback type. One learner may need explicit direct assistance and another indirect assistance, to self-correct. Similarly, simpler language structures may need less assistance than more complex ones (Aljaafreh & Lantolf, 1994).

2.3 Self-determination theory

For the sake of learner response to feedback, theories of learner motivation are of interest. Self-determination Theory (SDT) states that people are born with three basic psychological needs: to experience a sense of competence, relatedness, and autonomy (Ryan & Deci, 2017). All three are vital for aspects of individual and societal functioning. Applied to WCF, too many corrections may affect a learner's sense of competence, and the type and number of corrections may affect the student-teacher relationship if they are perceived as criticism. Indirect feedback, without explicit corrections, may contribute to a feeling of competence when a learner is successful, and the opposite, when a learner fails to understand. Likewise, the ability to independently revise and improve one's own texts may contribute to a sense of autonomy in addition to feelings of mastery.

From a learner-centered perspective, current research has established that motivation plays a significant role in how students perceive, use, and react to WCF, and that some learners engage more deeply than others (Ferris et al., 2013; Zhang & Hyland, 2018). Bitchener & Storch (2016) supports this notion, stating that: "it is the quality of the learners' engagement with the written CF that may help explain why feedback results in L2 development" (p. 7).

3. Literature

3.1 *WCF in education*

In the Norwegian educational context, feedback is considered a subcategory of assessment. It has been a key area in the development and implementation of the latest curriculum reforms (LK06 and LK20). Following ideas of the British Assessment Reform Group (ARG) and the work of Black et al. (2004), formative assessment (FA), also referred to as assessment for learning (AfL), has been implemented on a curricular policy level. Hattie and Timperley (2007) found that effective feedback answers questions of where learners are going, what level they are currently at, and where they need to go next (to reach their goal). Their meta-study is often referenced in discussions of CF as it resonates with the main purpose of FA: to aid and promote student development during the process of learning. The Norwegian Directorate for Education and Training (UDIR) clarifies that students have better learning outcomes:

- When they know what they should learn and what is expected of them
- When they receive feedback on the quality of their work
- When they receive advice on how to improve
- When they engage in their own learning through self-assessment (UDIR, 2010)

These four principles, echoing the findings of Hattie and Timperley, guide Norwegian language teachers' work with WCF. Recent studies, however, suggest that many Norwegian teachers struggle to implement the practice of AfL in everyday work with feedback and assessment (OECD, 2011; Saliu-Abdulahi, 2019; Vattøy, 2020).

3.2 *Teacher practice*

Evans et al. (2010) asked L2 practitioners (1,052) from 69 different countries about their written feedback practice. Ninety-two percent revealed that they provide WCF because: it helps students notice/build self-editing skills/understand errors (448); students expect it (223); students need it to be understood (193); because language matters (72) and because feedback is the teacher's responsibility (51). The remaining 8% of respondents said that they did not correct errors because; they believe content, organization, and rhetoric are more important than linguistic accuracy; students should take care of grammar errors by themselves; error correction is ineffective; and because they do not want to overwhelm, threaten, or discourage their students (pp. 58–61).

Several studies have, however, revealed discrepancies between what teachers want to do, say they do, and what they actually do (e.g., Lee, 2008a, 2008b; Mao & Crosthwaite, 2019). Lee (2008b) found that Hong Kong secondary teachers (26) did not always practice according to their own beliefs. They felt constrained by an overall exam-oriented system as well as by parents' and the school administration's expectations of a detailed error-focused practice (p. 81). In a more recent Chinese study of college teachers (5), misalignments between teacher beliefs and teacher practice were connected to teachers' workload or large class sizes (Mao & Crosthwaite, 2019). The concept of teacher cognition, more specifically, how teachers' thoughts, knowledge and beliefs influence their actions and decision-making, can help explain teachers' feedback practice (Borg, 2006).

Research conducted in Norwegian EFL classrooms point in different directions. In a study based on classroom observations and interviews, Saliu-Abdulahi et al. (2017) found that most teachers organize single-draft writing with limited opportunities for further work with feedback and/or text revision. The dominant feedback pattern is to provide written comments in Norwegian, both in-text and as endnotes. Teachers in the study report that they encourage students to work with the feedback they receive but have little time for revision within regular hours of instruction. In contrast, Horverak (2015) found that the participating upper secondary teachers provide feedback before and after text revisions, in line with approaches for process writing. Burner (2015) studied lower secondary teacher and student perceptions of FA in EFL writing and found that students appreciate constructive feedback, preferably oral, and text revision. They do, however, find it challenging when feedback is too focused on errors, which they perceive as negative, or when they do not understand the content of the feedback (p. 635). In the same study, teachers mention lack of time as a key factor for not working on revision during school hours (p. 641). Vattøy (2020) gathered interview data from lower secondary teachers (10) and found that the push for accountability in the form of examinations, testing, and marking, is a challenge to the implementation of AfL principles for feedback provision (p. 6).

Both the current curriculum, LK20, and international research stress that students need to be active in self-editing as well as interact with their teacher and the WCF to benefit from- and enhance its effectiveness (Lee, 2020; UDIR, 2010). This suggests that an important part of teachers' WCF practice is to facilitate time for work with revision, together with the teaching of revision strategies, so that students get the opportunity to process the feedback and develop self-editing skills.

3.3 Different views on the effect of WCF

In language learning research, two disciplines have taken a particular interest in the efficacy of WCF, but from slightly different angles. L2 writing research investigates the role that WCF may play in learners' development of editing and revision skills and their overall improvement in writing, also referred to as the learning-to-write dimension (Ferris, 2006). SLA research, in contrast, is more focused on the writing-to-learn perspective, i.e., whether WCF can facilitate the long-term acquisition of different linguistic features (e.g., Bitchener & Knoch, 2009; Sheen, 2007). As a result, separate theoretical and methodological traditions have developed, which complicates discussions of feedback efficacy. James Truscott (1996) famously claimed that feedback on grammar is harmful to students writing and that successful revision post-feedback only proves successful revision, not learning. Proof of learning would require successful independent use in new pieces of written production. This sparked a decade-long debate between Truscott and Ferris (e.g., Ferris, 1999, 2004, 2006; Truscott, 1996). Their publications attracted vast scholarly attention and contributed to considerable progress in the field.

In answer to some of Truscott's critique, a series of SLA studies on singular language items, such as definite and indefinite articles (Bitchener & Knoch, 2008, 2009, 2010; Ellis et al. 2006), and weak verbs (Frear & Chiu, 2015), applied a pre-test/post-test/delayed post-test design, with the use of control groups. They report clear and consistent advantages for students who receive written CF over control groups that receive no written CF. Different variables of written CF have also been examined, such as whether the feedback provided is focused or unfocused (Bitchener et al., 2005; Ellis et al., 2008), direct or indirect (Ferris, 2010), or provided together with error codes or in-class instruction (Ellis et al., 2008).

A more recent debate in terms of feedback efficacy concerns the applicability of focused CF research to real teacher practice. Van Beuningen et al. (2012) claim that studies that have focused on one or a few linguistic items only, hold little ecological value because a comprehensive or unfocused approach to CF provision is closer to the correction methods used by teachers. Teachers correct student texts to improve accuracy in general, not just one grammatical feature (Lee, 2020; Ferris, 2010; Storch, 2010). Moreover, Bruton (2009) raises concerns about focused CF studies because the WCF may be perceived as grammar exercises rather than authentic writing tasks. If this makes students monitor the target feature more consciously, the evaluation of the feedback effect becomes less representative of the actual effect.

A few studies have made attempts to increase ecological validity in CF research, looking specifically at comprehensive WCF to all or most language errors (Bonilla López et al., 2017; Truscott & Hsu, 2008; Van Beuningen et al., 2012). Truscott & Hsu (2008) investigated 47

university-level ESL learners. They provided comprehensive feedback to half the group and no feedback to the other half. The study found that even though the feedback improved students' accuracy in revision, there was no learning effect from the WCF in subsequent learner production. Bruton (2009) pointed out that this could be due to a "ceiling effect"; that is, the participating students made such few errors in the first place that measuring improvement became difficult. Bruton then performed a detailed re-analysis of Truscott & Hsu's data and found evidence of a "carry-over effect" from the first corrected text to the subsequent new text. This was hidden in the original report as Truscott and Hsu compared global error scores between the CF group and the control group; their calculations did not differentiate between repeated and new errors. Bruton maintains that this type of comparison contributes to false ideas of the CF effect and language gains (p. 139). His finding accentuates the need for qualitative studies of the CF effect in subsequent pieces of writing, preferably longitudinal.

4. Method

4.1 Design

The present study was designed as a longitudinal explorative case study involving three students and their two teachers. An error- and feedback analysis was considered most suitable to obtain an in-depth understanding of the teachers' WCF practice in its primary form. During work with the error and feedback analysis, questions about the teachers' practice started to emerge, and both teachers were invited to participate in a semi-structured interview to include their views and perspectives on WCF.

The study contains five repeated observations of between 600 to 1300 words in each text for each of the three students. The total number of words for error analysis is approximately 14,000. In addition to the 15 original texts, 9 revised versions were collected. The student texts were written during obligatory mock exams following the timeline in Table 1:

Table 1: Timeline for mock exams

School year	Month of collection	Mock Exam with teacher WCF	Revised version
Year 8 Fall		Not available	Not available
Year 8 Spring	May	Collected	Not available
Year 9 Fall	November	Collected	Collected
Year 9 Spring	April	Collected	Not available
Year 10 Fall	November	Collected	Collected
Year 10 Spring	May	Collected	Collected

4.2 Participants

The participating teachers and students are part of a corpus linguistic research project, TRAWL (Tracking Written Learner Language) (Dirdal et al., 2022), which enabled the longitudinal approach. Texts had already been collected for three years, at the onset of the study. The students had Teacher 1 in grades 8 and 9, and Teacher 2 in grade 10. The two teachers, however, have different educational backgrounds: Teacher 1, an MA in English and with several years of experience teaching English, whereas Teacher 2 has 20+ years of general teaching experience, and seven years of teaching English, but no formal English degree.

To get a clear idea of the teachers' WCF practice, it was essential to select a group of students that was as homogenous as possible in terms of L1 background, exposure to English, and level of proficiency. At the same time, their written production needed to be on a level that required WCF. The selected participants, two girls and one boy, all have Norwegian language backgrounds. They were holistically rated either grade 4, or slightly below 4 on all five mock exams by their teachers. Norway uses a 1–6 scale for grading, with six being the top mark. A quick analysis of the students' first 500 words of the first mock exam using the Textinspector online tool for lexical profiling² placed all three approximately between the B1 and the B2 level of the CEFR scale³.

4.3 Error and feedback analysis

A five-step error analysis was applied (Ellis & Barkhuizen, 2005, p. 57):

1. Collection of a sample of learner language
2. Identification of errors
3. Description of errors
4. Explanation of errors
5. Evaluation of errors

To keep the scope of the study manageable, steps 4 and 5 (explanation and evaluation of the errors) had to be kept at a minimum.

As previously mentioned, the study does not separate between errors and mistakes, the exception being clear slips on the keyboard (i.e., **søn* for son). Ambiguous cases are marked for

² <https://textinspector.com/>

³ <https://www.coe.int/en/web/common-european-framework-reference-languages/table-1-cefr-3.3-common-reference-levels-global-scale>

more than one category but only count as their most likely error type. Repeated errors of the same type that appear in the same text are counted separately, which will have an impact on tables showing general overviews of errors and feedback.

4.4 Categorizing errors and WCF

Ellis and Barkhuizen (2005) suggest a taxonomy of either linguistic- or surface structure categories, or a combination of both (p. 60). An error taxonomy was developed based on the Eckstein & Ferris (2017) linguistic categories and the Surface Structure Taxonomy (Dulay, Burt, & Krashen, 1982). It has four main categories (mechanical, grammatical, structure, and word choice), forty subcategories, five surface structure categories, and explanatory categories for word choice and spelling. See appendix 1 for the complete table. The combined taxonomy provided a detailed picture of the error focus of the teachers' WCF.

To track WCF alongside learner errors, the teachers' WCF was categorized and marked in the same spreadsheet as the errors. In accordance with Ellis' taxonomy of feedback types (2008b), the two variables, direct and indirect, were used. Any comments alongside the WCF, typically in the margin, including metalinguistic clues or explanations, were noted in a teacher comment section, and later included in the analysis of the teachers' feedback practice.

4.5 Data analysis

The learner errors and teacher WCF were analyzed in steps. First, each learner's text was read without the teacher's WCF. Secondly, the text was marked for all errors and a suggested corrected version was added in parentheses. In the next step, every error was classified, categorized, and explained with the taxonomy of error types. Once each student text had been error analyzed, it was compared to the version with teacher WCF. A tracking system for both errors and WCF was created. Notes were taken throughout the process as to whether students had attempted to successfully or unsuccessfully revise the WCF points. The figure below is an example of the spreadsheet.

Figure 1: Example of the spreadsheet with categorizations

Error id	Error	in context	Errortype	Category	Suggestion for corre	Explain	Norwegian
A1001	god	I am not a GOD driver	Error	M1.6	I am not a good driv	Amb	god
A1002	taken	I don't have TAKEN	Error	WC3.14	I don't have	L1	jeg har ikke
A1003	(a)	(∅) driver licence	Error	N3.1	a driver licence	Amb	

To answer research question 2, errors in the Mechanical subcategory *Spelling* and the Grammatical subcategory *Verbs* were tracked from one text to the next. The categories were chosen because they had both received one or more feedback points and been revised by the students, which increased the probability that students had noticed the WCF. The spelling errors were counted and compared to the number of tokens in each text using the formula (number of errors/number of tokens x 100). The selected verb errors were compared to the obligatory occasions for verb use to evaluate whether this would change the impression of improvement (number of errors/obligatory occasions for use x 100). For more about obligatory occasion analysis, see Ellis & Barkhuizen (2005, pp. 73–92).

4.6 Semi-structured interviews

The analysis of teacher WCF generated questions that called for the teachers' perspectives. A semi-structured interview guide was created (see Appendix 2). Teacher 2 agreed to a face-to-face interview conducted in Norwegian. It was recorded, transcribed verbatim, and translated into English. Teacher 1 answered the same questions via email in English, which was followed up by a brief conversation for clarifications. A general inductive approach (Thomas, 2006) was applied to condense the raw data into a summary format. The content was sorted according to the different themes of the questions since the teachers' answers sometimes overlapped from one theme to the other.

4.7 Notes regarding reliability

Researcher bias depends on both internal and external variables (Dörnyei, 2007, p. 60). For an error analysis of close to 1,400 errors combined with a feedback analysis, it is easy to make mistakes or overlook details due to reading fatigue. The comprehensive approach to the error analysis may further provide inconsistencies in the error categorization. However, the five different points of data collection, together with the amount of text and feedback, provide a 'thick' description of learner language and teacher WCF (Dörnyei, 2007). This, together with the longitudinal authentic nature of the data, makes the study valuable, albeit not generalizable. Regarding the participating teachers, it is important to mention that the collected WCF data shows but a fraction of their overall feedback practice, which comprises various kinds of written CF for different types of writing, as well as oral feedback in the classroom and the occasional teacher-student conference.

5. Results

5.1 Learner Error Context

In terms of the learner error contexts, the number of errors vary both between students and texts. All three students struggle with incomplete sentence structure and wordiness, which sometimes affects the comprehensibility of their writing. The subcategory *Verb* is another area of difficulty. The students have problems with the inflection of regular and irregular verbs, mainly in the past perfect and the simple past, but they also mix past and present tense in the same sentences. Additionally, many errors were found in the category *Word Choice*, and the subcategory *Spelling*.

As expected, the longest texts have the highest number of errors. When calculating global error scores (number of errors/number of tokens x 100), without considering recurring errors vs new errors or development within the various categories, the students' texts come across like this:

Table 2: Global error scores

	Exam 1	Exam 2	Exam 3	Exam 4	Exam 5
Student A	9,30 %	8,00 %	7,40 %	7,30 %	7,20 %
Student B	6,00 %	5,60 %	6,10 %	6,40 %	5,90 %
Student C	7,20 %	5,80 %	5,90 %	6,00 %	4,10 %

The percentages above suggest that Student A and C improved in accuracy from the first exam to the last, while Student B improved less. It should be noted that Exam 5 was Student C's second longest text (1022 tokens) and still the text with the lowest error rate. For examples of learner errors and detailed presentations of each learner profile, see Berg (2020).

5.2 Teacher WCF practice

At the time of the study, the school had participated in a national project to implement the new regulations of AfL (see section 3.1). The teachers were encouraged to provide WCF during the students' writing process rather than post-production. The collected data, mock exams, are typically written during a limited amount of time (one day only). Even though this may limit process writing approaches, the texts are also some of the most extensive examples of written production one can collect from a classroom context. Notably, Google Classroom was introduced as a platform in grade 10. As a result, Teacher 2 provided WCF digitally both during and after the texts were produced.

5.3 Teacher 1

5.3.1 Observed and self-reported practice in grade 8 and 9

Teacher 1 reports that regarding WCF, her general approach is that content is the most important, followed by structure, then errors on a sentence level, then grammar errors, word errors and last spelling. This does not mean that she ignores spelling errors, but they need to affect communication for her to point them out. She does, however, say that she often highlights words that are misspelled so that students can ‘fix it’ before handing in a final draft. For end-of-text comments, Teacher 1 explains that she mainly focuses on what the students achieve, and that she notes a few points on areas of improvement.

When asked about timing, she explains that for process writing, she provides most of the feedback on the first drafts:

I try to enter the text as early as possible into the process. I want to cheer and point out what they do well right away – to give them a boost and make them feel good about writing.

Teacher 1 also admits that despite her efforts to focus her WCF on the first drafts, she sometimes ends up commenting on finished texts:

I try to only give a comment on the final draft, but I sometimes end up correcting and highlighting on the last draft too – especially if the student appreciates it and I think he or she will learn from it.

Table 3: Teacher 1 approach to WCF, grades 8 and 9

Student A	Exam 1	Exam 2	Exam 3
Amount of WCF	50 WCF points, 150 errors total	9 WCF points, 89 errors in total	21 WCF points, 122 errors in total
Focus	22 different error types	6 different error types	12 different error types
WCF Type	mainly indirect	indirect	mainly indirect
Timing	after finished product	after finished product	after finished product
Student B			
Amount of WCF	21 WCF points, 77 errors in total	17 WCF points, 96 errors in total	48 WCF points, 178 errors in total
Focus	13 different error types	15 different error types	21 different error types
WCF Type	mainly indirect	direct and indirect	mainly indirect
Timing	after finished product	after finished product	after finished product
Student C			
Amount of WCF	52 WCF points, 114 errors in total	19 WCF points, 80 errors in total	36 WCF points, 82 errors in total
Focus	26 different error types	14 different error types	19 different error types
WCF Type	60/40 indirect/direct	indirect/direct	indirect/direct
Timing	after finished product	after finished product	after finished product

Table 3 presents a detailed record of Teacher 1's approach to WCF in the collected data: WCF focus, WCF type, and timing for providing WCF.

Table 3 shows that all feedback is provided after the finished product has been handed in, which is expected considering the time limitations of mock exams. It also describes a practice that varies between students and texts, mainly with the use of indirect WCF. When asked about whether she prefers to use direct or indirect WCF, Teacher 1 says:

I try to point out the errors first, so that the students can figure out what to do themselves. If they don't manage on their own, I try to help them.

Adding more detail to the teacher's feedback focus, Table 4 depicts the total amount of WCF according to error category.

Table 3: The total amount of WCF to error categories in grade 8 and 9

Student A				Student B				Student C			
Category	WCF	Indirect	Direct	Category	WCF	Indirect	Direct	Category	WCF	Indirect	Direct
Mechanics	25	25		Mechanics	19	18	1	Verbs	21	15	6
Verbs	23	22	1	Word Choice	16	10	6	Mechanics	20	16	4
Word Choice	18	14	4	Sentence Struct	15	9	6	Pronouns	12	11	1
Nouns	7	5	2	Verbs	10	8	2	SV-agreement	12	12	
Pronouns	3	2	1	Nouns	6	4	2	Word Choice	12	5	7
SV-agreement	2	2		Punctuation	4	4		Sentence Struct	9	6	3
Punctuation	1	1		Prepositions	4	3	1	Nouns	7	2	5
Sentence Struc	1	1		SV-agreement	4	4		Punctuation	8	3	5
				Adjectives	3	3		Prepositions	4	2	2
				Pronouns	3	2	1	Adjectives	1		1
				Adverbs	2	1	1	Adverbs	1	1	
	80	72	8		86	66	20		107	73	34

Looking at Table 4, most error categories have been targeted, however, the majority of feedback points are found in Mechanics (spelling), Verbs, Word Choice, and Sentence Structure. It becomes clear that Student A receives less feedback to fewer types of errors and with less direct WCF compared to Student B and C, even though she made the most errors out of the three. The numbers indicate that Teacher 1 adjusts the amount, focus and type of WCF to each student individually, but not according to the number of errors made. Considering how Teacher 1 claims to consider word errors and spelling errors last, it is interesting that the feedback points to Word Choice and Mechanics collectively amount to 40 % of all WCF points she provides. In terms of feedback type, the Word Choice category is targeted with the highest numbers of direct WCF in contrast to Mechanics with the highest numbers of indirect WCF.

5.3.2 Amount of feedback

The total amount of WCF points suggests that Teacher 1's feedback practice is somewhere between semi-focused and comprehensive, given that the definition of comprehensive WCF is to provide feedback to most or all errors in a text.

She explains that there are limitations to the amount of WCF one may provide:

Error correction may be experienced as criticism, and they can only take so much.

A clever student who is ambitious – I try to give as much feedback as possible. A student who struggles is only given what is necessary to be able to communicate.

When asked about her beliefs concerning WCF and how effective error correction is to help improve students' writing, Teacher 1 says:

I am not sure. I don't think it is very motivating to only focus on errors. Then again... If the student is really motivated, perhaps it works?

And

I don't believe giving the correct spelling gives them good strategies for learning. Sometimes, with some students, it feels right, though. They are so different – and you must use your gut feeling when it comes to these things.

5.4 Teacher 2

5.4.1 Observed and self-reported practice in grade 10

Teacher 2 provides fewer WCF points, more direct WCF than Teacher 1, and sometimes comments in Norwegian. He explains that he adapts the feedback to the level of the student in terms of explicitness (direct or indirect WCF):

I work differently with different students. I explain more to weaker students, but for the higher-level ones I may just write a question mark or NB! In the margin.

When asked about his language use in comments, Teacher 2 explains that it is important that students understand what he tries to tell them, and that for higher-level students he may comment in English. When elaborating on the matter, he expresses concern for the weaker students. He explains that the choices he makes relate to making students who struggle with English feel safe. He often tells them that it is ok to make errors.

In terms of timing, Teacher 2 says that it largely depends on how much time they have at hand. During a recent period of homeschooling, he had more time to work with process writing and to give feedback. He started using writing prompts only looking at content in one text and then specific language features in another. Otherwise, he provides feedback both as the students are writing, in shared Google documents, and after the students' writing has been graded:

But I tend to correct on the last version too, even though I know that they are most likely not going to revise it... I do it so that those who want to revise can do so."

I also experience that if I do not mark errors, the students don't understand the grade they get if the text looks error-free.

Table 5 provides a more detailed record of Teacher 2's approach to WCF in the collected data: WCF focus, WCF type, and time for providing WCF.

Table 4: Teacher 2 approach to WCF, grade 10

Student A	Exam 4	Exam 5
Amount of WCF	7 WCF points, 56 errors in total	5 WCF points, 63 errors in total
Focus	6 different error types	4 different error types
WCF Type	mainly indirect	direct and indirect
Timing	during writing and after	during writing and after
Student B		
Amount of WCF	5 WCF points, 44 errors in total	32 WCF points, 73 errors in total
Focus	4 different error types	13 different error types
WCF Type	direct	mainly direct (except for one category)
Timing	during writing and after	during writing and after
Student C		
Amount of WCF	30 WCF points, 78 errors in total	20 WCF points, 77 errors in total
Focus	9 different error types	10 different types
WCF Type	mainly direct	mainly direct
Timing	during writing and after	during writing and after

Table 5 shows that Teacher 2 provides WCF during the exam and after in end-of-text comments accompanying the assessment and grade. The number of different error types and the total amount of WCF points differ between students, texts, and compared to the approach of Teacher 1. When describing his approach to WCF, he says:

When we did corrections by hand, I sometimes used error codes, but it gets too complicated to do it digitally ... so now if I see the same error many places, I may mark all of them but only comment on one. The way I correct, I insert comments on the side (in the margin), and I write and explain... there=der, or there/their.

Teacher 2’s answer above explains the higher numbers of WCF points to Student B (Exam 5) and Student C (Exam 4), where he had marked or corrected all of one to two types of spelling errors.

Table 5: Feedback focus by error category

Student A				Student B				Student C			
Category	WCF	Indirect	Direct	Category	WCF	Indirect	Direct	Category	WCF	Indirect	Direct
Mechanics	7	4	3	Mechanics	21	15	6	Mechanics	28	7	21
SV-Agreement	3	2	1	Sentence Struct	5	1	4	Sentence Struct	7	2	5
Punctuation	1	1		Punctuation	2		2	SV-Agreement	4	1	3
Sentence Struct	1	1		Prepositions	2		2	Verbs	4	1	3
Word Choice	1	1		SV-Agreement	2		2	Word Choice	3	2	1
				Word Choice	2	1	1	Nouns	2		2
				Adjectives	1		1	Prepositions	2		2
	13	9	4		35	17	18		50	13	37

Table 6 further shows that Teacher 2 targets such few error types that it comes close to a focused approach, however, it is difficult to say whether this is a sign of a conscious strategy. When asked whether he selects a few specific errors for WCF focus, he says:

No, I can’t select like that ... but on a few occasions ... on short texts ... I have told them I will only look at verbs.

5.4.2 Amount of feedback

Looking at Table 5 and Table 6 together, Teacher 2 provides rather small amounts of WCF to the three students’ texts. However, when comparing the feedback points between texts and students, there is variation in focus and amount that speaks for an individually differentiated practice. Teacher 2 expresses concern for what the students prefer and how they may react to the amount of WCF:

It depends on the student. If I know that this is a student that wants to have all the errors corrected, then I correct more. But then it depends on the number of errors ... I don’t want it to look too bad when the students get their paper back.

Teacher 2 believes that one of the biggest challenges to teaching English, and to the possible effect of WCF, is the difference between students of low- and high-proficiency:

It's effective for those who already know... The ones you try to help, are the most difficult to help... and the ones who need your help the least are the ones who benefit the most.

And when talking more about how WCF may help students improve, Teacher 2 mentions aspects of motivation:

Ultimately with feedback, the students who try to understand the feedback, who make use of it, they have the energy and will to improve ... they are easy to help. But how can you make the other students get to that same point? I think you have to make it interesting ... and that thing with errors ... you have to make them understand that it is ok to try and to fail. That an error is not such a big thing.

Considering that the collected student texts and WCF data only follow three students, the interview data provided essential information about the teachers' practice, in general, as well as related to the participating students.

5.5 Tracking errors and WCF in subsequent texts

RQ2 was created to see whether it was possible to track any positive effects of WCF, from one text to another, to the next. Because the collected data stretches over three years of instruction, many factors, other than WCF, could impact visible improvement in the students' written production. Therefore, only error categories that had been targeted by the teachers' WCF, and been revised by the students, were chosen for analysis.

In Table 7, the frequency of errors in spelling (M1.6–M1.9), excluding issues with capitalization, is calculated against the total number of tokens in each text. The starting point was Exam 2, as this was the first exam where the students could revise and turn in a second version of their text.

Table 6: Error frequencies for the subcategory spelling (M1) in the category of Mechanics (M)

Mechanics	Exam 2				Exam 3				Exam 4				Exam 5			
	Error	Tokens	%	WCF Rev	Error	Tokens	%	WCF Rev	Error	Tokens	%	WCF Rev	Error	Tokens	%	WCF Rev
Student A																
M1.6	6			2 1	3			1 no	9			2 3	3			1 3
M1.7	1			1 0	3			no	1			1 1	1			1 1
M1.8					1			no					3			2 3
M1.9					9			no								
Total	7	780	0,9	3 1	17	1129	1,5		10	615	1,6		7	658	1,1	
Student B																
M1.6	3			0	4			2 no	2			1 1				
M1.7					2			2 no								
M1.8	4			1 1	9			5 no	1				2			2 0
M1.9	1			0	3			no	1							
Total	8	1222	0,7		18	1573	1,14		4	545	0,7		2	622	0,3	
Student C																
M1.6	3			0	7			4 no	5			3 2	14			6 6
M1.7								no					1			
M1.8	1			1 1	2			2 no	10			7 5				
M1.9								no	3			2 2				
Total	4	861	0,5		9	966	0,93		18	693	2,5		15	1022	1,5	

E = total amount of errors, Tokens = number of words in the text, % = errors in relation to tokens, WCF = both direct/indirect, Rev = Revised

Three things about Table 7 are worth mentioning. First, the students only revise errors that have been pointed out to them, suggesting that they have yet to develop independent self-editing skills. Second, with small numbers, the question of how to count the errors will have an impact on the measurements. This aspect becomes clear when working qualitatively with a smaller data set but would have been difficult to spot in a larger study. In Table 7 above, all errors were counted even if an error appeared multiple times in the same text because it provides a more accurate picture of the teachers' WCF. This, however, changes the impression of improvement. Additionally, as the general accuracy score does not differentiate between new and repeated errors, it provides general ideas of improvement but fails to evaluate the actual feedback effect. Because the exams consisted of different tasks with vastly different topics, new errors appeared in each text, and previous topic-dependent errors naturally did not.

A second way to investigate the possible impact of WCF on accuracy is to compare erroneous use to successful use. An obligatory occasion analysis accounts for all the occasions where a language item or rule must be present (Ellis & Barkhuizen, 2005, p. 80). This study calculated error frequency in selected verb form categories against an analysis of obligatory occasion for the same.

Table 7: *Obligatory Occasion Analysis vs Errors*

Verbs	Exam 2					Exam 3					Exam 4					Exam 5				
	OC	Error	%	WCF	Rev	OC	Error	%	WCF	Rev	OC	Error	%	WCF	Rev	OC	Error	%	WCF	Rev
Student A																				
Present perfect																				
regular						1	0			no										
irregular	3	0				2	0			no	7	2			0					
Simple past																				
regular	22	2		1	1	35	0			no										
irregular	47	8		2	6	47	2		1	no	6	0				1	0			
Past perfect																				
regular	1	0				3	3		1	no										
irregular						7	6		6	no	1	0								
total	73	10		13,6		95	11		12		14	2		14,2		1	0			
Student B																				
Present perfect																				
regular	2	1		1	0	1				no	3					3				
irregular	5					1				no	4									
Simple past																				
regular	17					65				no	1					39	1			1
irregular	57	1			0	87	1			no	4					27				
Past perfect																				
regular	2	2			0	4	2			no										
irregular	2	1			0	4	3		1	no						1				
total	85	5		5,9		162	6		3,7		12	0				70	1		1,4	
Student C																				
Present perfect																				
regular	1	1								no	1					1				
irregular	3	1		1	1	1				no	4					2				
Simple past																				
regular	33	2			0	35	2		1	no	1					17				
irregular	67					54	1		1	no	5					40				
Past perfect																				
regular	1					4	4		2	no						3				
irregular	3	1		1	1	4				no						1	1			
total	108	5		4,6		98	7		7,1		11	0				64	1		1,5	

OC = Obligatory occasion, Error = all errors, % = number of errors/number of obligatory occasions x 100, WCF = all types, Rev = Revised

Table 8 shows how errors in the pre-selected verb categories change from Exam 2 to Exam 5. The table includes present perfect, simple past, and past perfect (in both passive and active form). All three tenses were found in the subcategories V1.3 and V2.3 (see appendix 1). When looking at the error count only, all three students make major improvements. However, Exam 4 differed from the other exams as the students chose to write most of it in present tense. This had a natural, and visible, impact on the tracking of improvement. In Exam 5, Students B and C returned to writing in past tense, whereas Student A did not. With the added numbers for obligatory occasions, both Student B and C show improvement. Understanding how much the students improved and how few errors they made in the first place was possible through the comparison of right vs erroneous use of the verb forms in question. The analysis in Table 8 further supports that it was impossible to draw any conclusions about Student A's potential improvements without the obligatory occasion analysis, seeing how there was only one obligatory occasion in Exam 5 that required the past tense.

Even though the collected data provides a detailed picture of feedback provided to the students, it only shows a limited view of the teachers' practice, especially seeing that Teacher 2 only worked with the students one out of three years. There is no knowing whether Teachers 1 and 2 used a different approach to other students of lower or higher proficiency, however, their self-reported practice suggests so. The interview data thus added information that was necessary to better understand how the teachers think and work with WCF.

6. Discussion

Many scholars view authentic teacher WCF as predominantly concerned with students' overall improvement in writing (Lee, 2020; Van Beuningen et al., 2012), which is why they describe teacher feedback practice as 'comprehensive', but it is perhaps the intention that is comprehensive, not the amount of feedback.

In this study, the many error types targeted with WCF indicate a comprehensive approach, but the varying and sometimes scarce amounts of WCF do not. Even though the two teachers work with the same students, their approach to providing WCF differs. Parts of the variation may be explained by the different writing tasks or that the students had developed over time, while some are clearly related to teacher beliefs: e.g., the teachers' contradictory views on how to treat spelling errors with either direct or indirect feedback, and visible differences in the number of feedback points.

Both data from the interviews and the collected student texts confirm that the participating teachers are guided by their evaluation of student needs, their own beliefs about how to use WCF to meet those needs, and external factors, such as time and the technical platform for writing. As such, the two lower secondary teachers' WCF practice can be explained with the concept of teacher cognition (Borg, 2006). Misalignment between beliefs and practice was found in how the teachers reported providing WCF to finished or graded texts, knowing that the feedback possibly would go unnoticed. There were further discrepancies between the self-reported practice and actual practice of Teacher 1, in regard to word choice and spelling errors, but too little data to draw any conclusions.

The analysis of teacher WCF revealed that in terms of focus, certain local surface errors are provided with repeated corrections in the same text while most other errors are pointed out just once or twice. From a cognitive perspective, drawing attention to an error once may not be sufficient if competing input makes it less visible. Likewise, too much input may lead to information overload. Looking at the students' texts, there are examples where the students

seemingly notice (i.e., they revise), but continue to produce the same error post WCF and successful revision. One explanation is that more practice and processing is needed for consolidation to take place (Bitchener, 2019). Alternatively, the students simply copied the teacher's corrections without understanding or processing the information.

From a sociocultural perspective, repeated errors post-successful revision suggest that the learner needs continued direct assistance before being able to self-correct. The theory of ZPD states that for WCF to be efficient, it needs to enable learners to perform beyond their current capacity (Sheen, 2010). This aspect of SCT supports individual differentiation in WCF focus and type (indirect, direct, or metalinguistic), depending on the learner's proficiency. One learner may need more explicit instructions for one type of language error compared to other types of errors, and other types of learners (Aljaafreh & Lantolf, 1994). The teachers' observed and self-reported WCF practice shows evidence of such differentiation.

The question, then, is what should be considered the learners' ZPD in terms of language errors in writing? Are self-correctable spelling errors beyond the students' current capacity, and do they need the teacher's assistance to improve? Or should the teacher focus on an area where the student repeatedly produces the same type of error? Teacher 1's high numbers of indirect WCF to spelling errors indicate that she believes in her students' capacity to self-correct, which suggests that the scaffolding could be directed elsewhere. However, if the WCF was provided to help improve the overall impression of the text, rather than assist the learning process, the focus on self-correctable errors would be justified.

Seeing the collected data from a motivational perspective, it is striking how little the students engage in revision and the feedback they receive. At the same time, the teachers' reported beliefs about WCF suggest that many of the decisions they make are based on concerns about student motivation, either not wanting to overwhelm the students, or providing indirect feedback to let the students find solutions on their own. Because the students did not engage much with the collected WCF, it was difficult to a) discern any effect on their motivation, sense of competence, autonomy, or relatedness with their teacher, and b) investigate its effectiveness on their written accuracy as well as their learning process.

In the second part of the study, an attempt was made to track the positive WCF effect on accuracy in subsequent learner writing. The data suggests that the students improved, but the relation between accuracy gains and WCF was unclear due to the differences between the writing tasks. Instead, valuable insights were made related to issues with measurements of accuracy.

The process made it clear that it is easier to track rule-based errors compared to errors in spelling or word choice, perhaps explaining the focus of previous research (e.g., Bitchener &

Knoch, 2008, 2009, 2010; Ellis et al. 2006). Because words or expressions of low frequency often relate to topic, setting, and genre, and teachers ask students to practice various kinds of writing on multiple topics, misspelled words are less likely to reappear in subsequent texts. Therefore, global measurements of accuracy are more suitable for errors in word choice and spelling, especially for studies of written learner language produced in non-experimental settings. Because it does not differentiate between new and repeated errors (Bruton, 2009), it may, however, not be valid as an indicator of feedback efficacy.

The detailed re-reading of the students' texts, enabled by the case study format, revealed that decisions regarding what to include when defining and evaluating errors, as well as whether to count repeated errors or not, changed the impression of improvement. This insight has implications for how we interpret larger quantitative studies of the WCF effect, as it raises questions about the calculations. The added obligatory occasion analysis (Ellis & Barkhuizen, 2005) provided new perspectives on how to measure accuracy, as well as new understandings of the learner language already analyzed. It is a reminder that we need to observe errors and successful use to fully understand a student's level of competence.

7. Conclusion

The present study aimed to explore authentic longitudinal comprehensive teacher WCF, the participating teachers' beliefs about WCF, and possible improvements in accuracy following the teachers' WCF in subsequent writing. Its main limitations derive from the small number of participants, and that it did not control for student engagement with the WCF. This made it difficult to say whether improvements in accuracy were a result of the WCF provided, and whether patterns in the teachers' observed practice could be understood as representative of their overall WCF practice, or not. The qualitative and longitudinal approach did, however, provide insight into the complex nature of authentic learner language, implications for using such data in research, and factors affecting teachers' WCF decision-making.

Norwegian guidelines for feedback practice state that learners need to be involved in the work with assessment and feedback to achieve better learning outcomes, something which requires action, processing, and response to the WCF learners receive. This is supported by cognitive theories, sociocultural theory as well as Self-Determination theory, albeit for varied reasons. In the present study, the teachers reported that they worked with process writing, but also that they tended to provide WCF to finished graded texts, knowing that the feedback possibly would go unnoticed. The data further suggests that the students did not engage with the feedback enough to benefit from its learning potential.

The attempt to track learner errors and WCF comprehensively proved difficult but worthwhile as it revealed how global accuracy measurements might not present an accurate picture of the feedback effect. This raises questions of validity concerning previous research and proof of WCF efficacy. The limited scope of the study suggests that more longitudinal research using larger datasets is needed to better understand authentic teacher WCF and its effect on learner language development. Future studies of teacher practice would benefit from including the student perspective on why, how, and when they actively engage with WCF. A combination of quantitative and qualitative approaches to error and feedback analysis is recommended.

References

- Aljaafreh, A., & Lantolf, J. P. (1994). Negative Feedback as Regulation and Second Language Learning in the Zone of Proximal Development. *The Modern Language Journal*, 78(4), 465–483.
- Anderson, J. R. (1983). *The architecture of cognition*. Harvard University Press.
- Assessment Reform Group (ARG). (2002). *Assessment for Learning: 10 Principles*. https://assessmentreformgroup.files.wordpress.com/2012/01/10principles_english.pdf
- Berg, E. M. (2020). Written Corrective Feedback and the Development of L2 Learner Language: A longitudinal study of lower secondary EFL writing in Norway [MA Thesis]. University of Agder, Norway.
- Bitchener, J., Young, S., & Cameron, D. (2005). The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 14(3), 191–205.
- Bitchener, J. (2019). The Intersection between SLA and Feedback Research. In F. Hyland, & K. Hyland (Eds.), *Feedback in Second Language Writing: Contexts and Issues* (pp. 85–105). Cambridge University Press.
- Bitchener, J., & Ferris, D. R. (2012). *Written corrective feedback in second language acquisition and writing*. Routledge.
- Bitchener, J., & Knoch, U. (2008). The value of written corrective feedback for migrant and international students. *Language Teaching Research*, 12(3), 409–431.
- Bitchener, J., & Knoch, U. (2009). The Contribution of Written Corrective Feedback to Language Development: A Ten Month Investigation. *Applied Linguistics*, 31(2), 193–214.
- Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback. *Journal of Second Language Writing*, 19(4), 207–217.
- Bitchener, J., & Storch, N. (2016). *Written corrective feedback for L2 development*. Multilingual Matters.
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the Black Box: Assessment for Learning in the Classroom. *Phi Delta Kappa*, 86(1), 8–21.
- Bonilla López, M., Van Steendam, E., & Buyse, K. (2017). Comprehensive corrective

- feedback on low and high proficiency writers. *ITL – International Journal of Applied Linguistics*, 168(1), 91–128.
- Borg, S. (2006). *Teacher Cognition and Language Education*. Continuum.
- Bruton, A. (2009). Designing research into the effects of grammar correction in L2 writing: Not so straightforward. *Journal of Second Language Writing*, 18(2), 136–140.
- Burner, T. (2015). Formative assessment of writing in English as a foreign language. *Scandinavian Journal of Educational Research*, 60(6), 626–648.
- Corder, S. P. (1967). Significance of learners' errors. *Review of Applied Linguistics*, 5, 161–170.
- 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. <https://doi.org/10.46364/njltl.v10i2.1005>
- Dörnyei, Z. (2007). *Research methods in applied linguistics: quantitative, qualitative, and mixed methodologies*. Oxford University Press.
- Dulay, H., Burt, M., & Krashen, S. (1982). *Language two*. Oxford University Press.
- Eckstein, G., & Ferris, D. (2017). Comparing L1 and L2 Texts and Writers in First-Year Composition. *TESOL Quarterly*, 52(1), 137–162.
- Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition*, 28(2), 339–368.
- Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, 36(3), 353–371.
- Ellis, R. (2008a). *The study of second language acquisition* (2nd ed.). Oxford University Press.
- Ellis, R. (2008b). A typology of written corrective feedback types. *ELT Journal*, 63(2), 97–107.
- Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford University Press.
- Evans, N. W., Hartshorn, K. J., & Allen Tuioti, E. (2010). Written Corrective Feedback: The Practitioners' Perspective. *International Journal of English Studies*, 10(2), 47–77.
- Ferris, D. R., Liu, H., Sinha, A., & Senna, M. (2013). Written corrective feedback for individual L2 writers. *Journal of Second Language Writing*, 22(3), 307–329.
- Ferris, D. (1999). The case for grammar correction in L2 writing classes: A response to Truscott (1996). *Journal of Second Language Writing*, 8(1), 1–11.
- Ferris, D. R. (2004). The “Grammar Correction” Debate in L2 Writing: Where are we, and where do we go from here? (and what do we do in the meantime ...?). *Journal of Second Language Writing*, 13(1), 49–62.
- Ferris, D. R. (2006). Does error feedback help student writers? New evidence on the short- and long-term effects of written error correction. In K. Hyland, & F. Hyland (Eds.), *Feedback in Second Language Writing: Contexts and Issues* (pp. 81–104). Cambridge University Press.

- Ferris, D. R. (2010). Second Language Writing Research and Written Corrective Feedback in SLA. *Studies in Second Language Acquisition*, 32(2), 181–201.
- Frear, D., & Chiu, Y. H. (2015). The effect of focused and unfocused indirect written corrective feedback on EFL learners' accuracy in new pieces of writing. *System*, 53, 24–34.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Horverak, M. O. (2015). Feedback practices in English in Norwegian upper secondary schools. *Nordic Journal of Modern Language Methodology*, 3(2).
- Krashen, S. D. (1984). *Writing*. Elsevier Science Limited.
- Lee, I. (2008a). Ten mismatches between teachers' beliefs and written feedback practice. *ELT Journal*, 63(1), 13–22.
- Lee, I. (2008b). Understanding teachers' written feedback practices in Hong Kong secondary classrooms. *Journal of Second Language Writing*, 17(2), 69–85.
- Lee, I. (2020). Utility of focused/comprehensive written corrective feedback research for authentic L2 writing classrooms. *Journal of Second Language Writing*, 49, 100734.
- Mao, S. S., & Crosthwaite, P. (2019). Investigating written corrective feedback: (Mis)alignment of teachers' beliefs and practice. *Journal of Second Language Writing*, 45, 46–60.
- McLaughlin, B. (1987). *Theories of second-language learning*. Arnold.
- McLaughlin, B. (1990). Restructuring. *Applied Linguistics*, 11(2), 113–128.
- OECD, (2011). Nusche, D., Earl, L., Maxwell, W., Shewbridge, C., & For, O. *OECD Reviews of Evaluation and Assessment in Education: Norway 2011*. OECD.
- Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Press.
- Saliu-Abdulah, D., Hellekjær, D., & Hertzberg, G. O. (2017). Teachers' (Formative) Feedback Practices in EFL Writing Classes in Norway. *Journal of Response to Writing*, 3(1), 31–55.
- Saliu-Abdulah, D. (2019). *Teacher and Student Perceptions of Current Feedback Practices in English Writing Instruction* [Doctoral thesis]. <https://www.duo.uio.no/handle/10852/66784?show=full>
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In *Attention and awareness in foreign language learning* (pp. 1–63). Honolulu: Second Language Teaching & Curriculum Center, University of Hawai'i At Manoa.
- Schmidt, R. (2001). Motivation and Strategy Use, and Pedagogical Preferences in Foreign Language Learning. In Z. Dörnyei & R. Schmidt (Eds.), *Motivation and Second Language Acquisition*. University of Hawaii's Press.
- Sheen, Y. (2007). The Effect of Focused Written Corrective Feedback and Language Aptitude on ESL Learners' Acquisition of Articles. *TESOL Quarterly*, 41(2), 255–283.
- Sheen, Y. (2010). Introduction. *Studies in Second Language Acquisition*, 32(2), 169–179.
- Storch, N. (2010). Critical Feedback on Written Corrective Feedback Research. *International*

Journal of English Studies, 10(2), 29.

- Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics: Studies in honor of H. G. Widdowson* (pp. 125–144). Oxford University Press.
- The Norwegian Directorate for Education and Training (UDIR). (2010). *Nasjonal satsing på vurdering for læring* [National incentive for assessment for learning]. <http://www.udir.no/Vurdering-for-laring/Nasjonal-satsing1/Nasjonal-satsing-pa-Vurdering-for-laring>
- Thomas, D. R. (2006). A General Inductive Approach for Analyzing Qualitative Evaluation Data. *American Journal of Evaluation*, 27(2), 237–246. <https://doi.org/10.1177/1098214005283748>
- Truscott, J. (1996). The Case Against Grammar Correction in L2 Writing Classes. *Language Learning*, 46(2), 327–369.
- Truscott, J., & Hsu, A. Y. (2008). Error correction, revision, and learning. *Journal of Second Language Writing*, 17(4), 292–305.
- Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2012). Evidence on the Effectiveness of Comprehensive Error Correction in Second Language Writing. *Language Learning*, 62(1), 1–41.
- Vattøy, K. D. (2020). Teachers' beliefs about feedback practice as related to student self-regulation, self-efficacy, and language skills in teaching English as a foreign language. *Studies in Educational Evaluation*, 64, 100828.
- Villamil, O. S., & de Guerrero, M. C. M. (2006). Sociocultural theory: A framework for understanding the social-cognitive dimension of peer-feedback. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 23–41). Cambridge University Press.
- Vygotsky, L. S., Cole, M., John-Steiner, V., Scribner, S., & Souberman, E. (1978). *Mind in Society: The Development of Higher Psychological Processes* (Revised ed.). Harvard University Press.
- Vygotsky, L. S. (1981). *Thought and language*. M.I.T Press.
- Zhang, Z. V., & Hyland, K. (2018). Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing*, 36, 90–102.

Appendix 1

Level 1: Main Categories (4)	Level 2: Subcategories (40)
Mechanical	M1. Spelling, M2. Capitalization, M3. Hyphens P1. Punctuation
Grammatical	
Nouns	N1. Plural Marker, N2. Possessive Marker, N3. Articles/Determiners, N4. NP
Pronoun Use	PU1. Subjective, PU2. Objective, PU3. Unclear Reference, PU4. Other
Adjectives	AA1. Adjective, AA2. Comparative, AA3. Adjective for Adverb
Adverbs	AD1. Adverb, AD2. Adverb for Adjective
Prepositions	PP1. Preposition
SV-agreement	SV1. Error in the NP, SV2. Error in the VP
Verbs	V1. Tense/Aspect Regular Verb, V2. Tense/Aspect Irregular Verb, V3. 3 rd Person Present Tense -s, V4. Passive Voice, V5. Progressive -ing, V6. Modal Auxiliary, V7. Mixing Past and Present, V8. Negation, V9. Interrogative
Sentence Structure	SS1. Run-on-sentence, SS2. Word order, SS3. Missing/unnecessary/wrong word, SS4. Comma splices, SS5. Relative clause, SS6. Other/subordinate that clause, SS7. Conjunction
Word Choice (Lexical Errors)	WC1. Lexical Transfer/L1 Influence, WC2. Lexical Intralingual/Developmental Errors, WC3. Word/Expression Made up (does not exist in L1 or L2), WC4. Other
Level 3	Explanatory
Surface Structure Taxonomy (for grammar and mechanical errors)	1. Omitted 2. Added 3. Misformed 4. Misordered 5. Other
Added Categories for Spelling	6. Incorrect 7. Phonological 8. Homophones 9. L1 influence
Categories for Word Choice	10. Direct Translation 11. False Friends 12. Preposition Error in a Phrasal Verb, L1 influence 13. Idiomatically incorrect 14. Idiomatically incorrect and direct translation of a L1 expression/word 15. Wrong Meaning for Context 16. Other

Appendix 2

Appendix 2				
Interview guide				
1. Briefly describe your educational background and experience as a teacher				
2. Describe your approach to teaching writing in your classroom (Do you provide writing prompts? How do you prepare students before writing tasks? Do you make students write multiple drafts? Do you use peer feedback? Do you use teacher-student writing conferences?)				
3. In your opinion, what kind of errors in a written text should be given immediate attention?				
4. Describe your approach to providing feedback to your students.				
5. What are the strategies you use to respond to students writing?				
6. Do you usually provide the students with the correct linguistic form? Yes – Why? No – Then what do you usually do? Do you underline/circle/locate the error in the text? Or do you comment in the margin that there is an error?				
7. Do you attempt to correct all the mistakes in the text? Yes – Why? No – Do you select specific errors? How and why?				
9. Do you use electronic feedback? (hyperlink)				
10. How do you respond to multiple drafts of student texts? To which draft do you usually provide error correction?				
11. To what extent do you rely on the rubric/knowledge requirement by the Directory of Education? Do you find it difficult to interpret the knowledge requirements?				
12. What is the most challenging aspect of providing feedback to students?				
13. In your opinion, how effective is error correction in improving student writing?				
14. How do contextual factors such as time, workload, <u>curriculum</u> and institutional policies affect your approach?				
15. How important it is for you to address the following in student texts?				
Type of errors	Not at all important	Slightly important	Moderately important	Very important
Mechanical errors				
Lexical errors				
Grammatical errors				
Organizational issues				
Content issues				