



Faculty of Education and Natural Sciences

Catharine Meissner

Master's thesis

**Formative assessment at the intersection of
principles, practice and perceptions**

Formativ vurdering i skjæringspunktet mellom prinsipper, praksis og persepsjoner

Lektorutdanning i engelsk

2018

Samtykker til tilgjengeliggjøring i digitalt arkiv Brage

JA NEI

Acknowledgements

When Charles Dickens wrote that it was the best of times and the worst of times, I think he was referring to writing a master's thesis. I am sure he was. This thesis is the culmination of five years of studies, five years of support, five years of friendship, five years of frustration, and many more things. There are so many people who deserve some credit for this thesis, that I cannot mention all, but here are a few.

For Bjørnar, thank you for cheering me on and always believing in me, and for taking care of everything these past six months. I look forward to spending time with you and Erle and Silas again.

For Stine, thank you for being the Sherlock to my Holmes (you know what I mean), and so much more throughout this process. I might have been able to do it without you, but I wouldn't have had as much fun. We have the best conversations.

For Torunn, thank you for being my friend, and for telling me when to get my act together. I have needed to hear that more than a few times these last six months.

For Elisabeth, Caroline, Stefanie, Mari, Matilde and Elise, the embodiment of smart, strong women, all role models in your own way. Thank you all.

For my excellent supervisors, Heidi Silje Moen and Christina Sandhaug. Thank you both for all the encouragement, critical and thoughtful comments, for making me think, and for helping me along when I was out of thoughts. I really could not have done this without you two.

Heidi, thank you for dealing with my countless emails at all hours of the day, all days of the week, and for all the time I know you have put into helping me.

Christina, thank you for those gold nuggets, the comments that pushed me along when I was stuck, you saved my bacon more than once.

Hamar, May 2018

Catharine Meissner

Table of Contents

Contents

ACKNOWLEDGEMENTS.....	3
TABLE OF CONTENTS.....	4
TABLE OF FIGURES.....	7
NORSK SAMMENDRAG.....	8
ENGELSK SAMMENDRAG (ABSTRACT).....	9
1. INTRODUCTION	10
1.1 THESIS AIM AND RESEARCH QUESTIONS	10
1.2 THESIS STRUCTURE	11
1.3 THE KNOWLEDGE PROMOTION REFORM-LK06.....	11
<i>1.3.1. Writing as a basic skill.....</i>	<i>12</i>
<i>1.3.2 Writing as a main subject area.....</i>	<i>13</i>
<i>1.3.3 The Assessment for Learning project.....</i>	<i>15</i>
1.4 DEFINING ‘ASSESSMENT’ AND OTHER TERMS	17
<i>1.4.1 Formative assessment and Assessment for Learning.....</i>	<i>17</i>
<i>1.4.2 Formative Assessment and Summative Assessment</i>	<i>18</i>
<i>1.4.3 Feedback</i>	<i>19</i>
2. THEORETICAL FRAMEWORK.....	21
2.1 PRINCIPLES OF FORMATIVE ASSESSMENT	21
2.2 FORMATIVE ASSESSMENT ISSUES	23
2.2.1 <i>Formative assessment in a behaviorist perspective.....</i>	<i>24</i>
2.2.2 <i>Formative assessment and cognitive theories.....</i>	<i>25</i>
2.2.3 <i>Formative assessment and social views of learning.....</i>	<i>26</i>
2.2.4 <i>Implications of theoretical grounding.....</i>	<i>27</i>

2.2.5	<i>Transparency, validity and reliability in formative assessment</i>	28
2.3	FORMATIVE ASSESSMENT AND MOTIVATION	31
2.4	THEORETICAL PERSPECTIVES ON WRITING	33
2.4.1	<i>The Response Triangle</i>	34
2.4.2	<i>The Wheel of Writing</i>	35
2.5	SUMMARY	38
3.	PREVIOUS RESEARCH	40
3.1	BEST PRACTICE RECOMMENDATIONS	40
3.2	WRITTEN FEEDBACK PRACTICES	41
3.2.1	<i>The focus of written feedback</i>	42
3.2.2	<i>How formative is the feedback?</i>	44
3.2.3	<i>Implications of feedback practices</i>	47
3.3	STUDENT ATTITUDES TOWARDS WRITTEN FEEDBACK	48
3.3.1	<i>Grades and feedback</i>	50
3.4	SUMMARY	52
4.	METHOD AND MATERIALS	53
4.1	PHASE 1-TEACHER COMMENTS	54
4.1.1	<i>Analysis of teacher comments</i>	55
4.1.2	<i>Strengths and limitations of using teacher comments</i>	56
4.2	PHASE 2- FOCUS GROUP INTERVIEW	58
4.2.1	<i>Analysis of the focus group material</i>	59
4.2.2	<i>Strengths and limitations of the focus group method</i>	60
4.3	ETHICAL CONSIDERATIONS	63
5.	FINDINGS	64
5.1	TEACHER COMMENTS	64

5.1.1	<i>Feedback focus</i>	64
5.1.2	<i>In-text comments</i>	65
5.1.3	<i>The formative potential of end comments</i>	67
5.2	TEACHER REFLECTIONS ON ASSESSING.....	68
5.3	FOCUS GROUP RESULTS	69
5.3.1	<i>Student preferences of teacher comments</i>	70
5.3.2	<i>Student perceptions of feedback practices</i>	73
5.4	SUMMARY AND SYNTHESIS	77
6.	DISCUSSION	79
6.1	THE PRIMACY OF LANGUAGE.....	79
6.2	CONTEXT AS A FACTOR IN ASSESSMENT	82
6.3	ASSESSMENT AS REASSURANCE	83
6.4	A COMPLICATED RELATIONSHIP	85
6.5	PREREQUISITES FOR ASSESSMENT QUALITY	87
7.	CONCLUDING REMARKS	91
	REFERENCE LIST	93
	APPENDICES	102
	APPENDIX 1- APPROVAL FROM NSD	102
	APPENDIX 2 INSTRUCTIONS FOR TEACHERS	103
	APPENDIX 3 SAMPLE OF STUDENT TEXT WITH INSERTED COMMENTS	104
	APPENDIX 4 INFORMATION FOR STUDENTS AND PARENTS	105
	APPENDIX 5 CATEGORIZED COMMENTS	106
	APPENDIX 6 CATEGORIZED COMMENTS	107
	APPENDIX 7 INTERVIEW GUIDE.....	108

Table of Figures

FIGURE 1 HOEL (2000), AFTER HILLOCKS (1987).....	34
FIGURE 2 THE WHEEL OF WRITING (BERGE, EVENSEN, THYGESEN, 2016)	36
FIGURE 3 THE WHEEL OF WRITING (BERGE, EVENSEN, THYGESEN, 2016)	37
FIGURE 4 DATA COLLECTION PROCESS.....	54
FIGURE 5 CATEGORY EXAMPLES.....	55
FIGURE 6 CATEGORIZED COMMENTS-SAMPLE.....	56
FIGURE 7 FOCUS GROUP CATEGORIES.....	60
FIGURE 8 TEACHER 2	66
FIGURE 9 TEACHER 3	66
FIGURE 10 TEACHER 4	66
FIGURE 11 TEACHER 1	67
FIGURE 12 TEACHER 6	68
FIGURE 13 TEACHER 4	68
FIGURE 14 TEACHER 5	68
FIGURE 15.....	70
FIGURE 16.....	71
FIGURE 17.....	72
FIGURE 18.....	72
TABLE 1 OVERVIEW OF TEACHER COMMENTS	65

Norsk sammendrag

Tidligere forskning på vurdering har rettet seg mot læreres vurderingspraksis, eller generelle forståelser av hva som kjennetegner effektive tilbakemeldinger. I tillegg har mye av denne forskningen blitt utført i førstespråkkontekster. Denne oppgaven er et bidrag til forskning på formativ vurdering av skriftlig engelsk som et andrespråk i Norge.

I denne oppgaven har jeg utforsket mulige utfordringer ved formativ vurdering ved å analysere tilbakemeldinger fra syv lærere på en elevtekst, og læreres tanker om hvorfor de valgte å kommentere som de gjorde. I tillegg gjennomførte jeg en fokusgruppe med fire elever på VG1, for å undersøke deres oppfatninger om tilbakemeldingspraksiser og -preferanser.

Analysen av lærernes tilbakemeldinger viser en tydelig tendens til å rette kommentarer mot lokale tekstnivåer, selv om det var store variasjoner i hvor spesifikke tilbakemeldingene var. Noen lærere påpekte at mangelen på kontekst var problematisk når de vurderte teksten, og følte at informasjon om eleven var nødvendig for å vurdere teksten.

Elevene i fokusgruppen hadde ingen tydelige preferanser for tilbakemeldingstype, annet enn at de ønsket «tydelige tilbakemeldinger som fikk dem til å tenke». Denne mangelen på klare preferanser var sannsynligvis et resultat av elevenes oppfatninger om vurderingspraksisene på skolen deres, fordi elevene ofte følte seg usikre på hvordan de hadde prestert.

Disse to perspektivene, lærernes tilbakemeldinger og elevenes oppfatninger, har bidratt til en større diskusjon i denne oppgaven, nemlig hva som er mulige problemer med formativ vurdering. Jeg diskuterer det problematiske forholdet mellom formative og summative vurdering, samt forestillingen om at konseptene validitet og reliabilitet er overførbare til hensikten og bruken av formativ vurdering.

Engelsk sammendrag (abstract)

Previous assessment research has focused on teacher assessment practice, or general understandings of what constitutes effective feedback. Furthermore, much of this research has been carried out in L1 contexts. This study is a contribution to research on formative assessment in written English as an L2 in Norway. In this study, I have explored possible challenges of formative assessment by studying feedback provided by seven teachers to a student text, and their thoughts on why they chose to comment as they did. In addition, a focus group was conducted with four students at grade VG1 in upper secondary, to elicit their perceptions of feedback practices and preferences.

The analysis of the teacher comments shows a strong tendency to provide feedback directed at local text levels, though the level of specificity among the different teachers was varied. Some teachers cited a lack of contextual clues as problematic when assessing, and felt that information about the student was necessary to assess the text.

The focus group students did not have a clear preference of comment types, though they stated that they wanted “specific comments that made them think”. This lack of strong preferences was likely due to their perception of the assessment practices at their school, as the students were often left feeling insecure of how they were performing.

These two perspectives, the teacher comments and student perceptions, have contributed to a wider discussion in this thesis, namely what the possible problems of formative assessment are. I discuss the problematic relationship between formative and summative assessment, and the notion that the psychometric concepts of validity and reliability are applicable to the purpose and use of formative assessment.

1. Introduction

This thesis is about the possible challenges of implementing theoretically sound formative assessment practices in assessment of writing in the English subject, and how students¹ perceive these practices. Writing is complex. It is about the processes that happen during the production of a text, but it is also about the finished product. Assessing writing means that the teacher needs to have in-depth knowledge of writing as process and product, as well as the distinctive characteristics of the subject the student is writing within. To help students develop their writing skills, it is necessary that teachers know what it means to be a competent writer, and that they can reflect on and use relevant theory to make informed teaching and assessment choices (Sandvik, 2011, p. 1). Central to assessment competence is knowing how views of learning impact the teaching and assessment choices one makes (p. 7), and how these choices affect fundamental principles of all assessment, namely transparency, reliability and validity. The problematic nature of these principles in formative assessment is a key issue that will be discussed in chapter 2. Assuming that feedback is instrumental to the process of becoming more adept at communicating in writing, providing quality feedback is one of the most important tasks of the teacher. However, what constitutes quality feedback is difficult to define, and is largely a matter of preference, though there is an abundance of response literature with ‘best practice’ recommendations on how to provide feedback (Straub, 1997; Hyland & Hyland, 2001; Lee, 2008; Ferris, 2014). This makes it all the more important to look at how teachers frame their responses, what they focus on, and how students perceive the response, in order to discuss the challenges of formative assessment.

1.1 Thesis aim and research questions

Most assessment research has focused on either teacher assessment practice, or what characterizes effective feedback in general. Consequently, there is a need for domain-specific research on formative assessment (Bennett, 2011, p. 15). Each subject has its own ‘culture of knowledge’, an understanding of what is important to know, how this is expressed through criteria, and what quality writing entails in different subjects (Evensen, 2009, p. 20). Recent studies on assessment practices in Norway have found that the degree to which feedback practices are formative in the English subject varies quite a lot (Burner, 2015a; Horverak, 2015, 2016; Saliu-Abdulah; 2017;

¹ I follow American writing conventions, and therefore use students, rather than pupils.

Saliu-Abdulah, Hellekjær & Hertzberg, 2017). These studies indicate a need for more research on how formative assessment is practiced in the English subject, and the challenges teachers face in trying to balance the needs of students with theoretical principles of formative assessment.

This thesis will explore the possible challenges of formative assessment at the intersection of assessment theory, teacher practice, and student perceptions by examining how texts by Norwegian L2 learners of English are assessed, and what types of feedback students find useful. The following research questions will serve to limit the scope of this thesis:

-What does formative assessment theory say about 'best practice'?

-What types of written feedback do teachers give students, and how do teachers explain their feedback practices?

-What are students' perceptions of written feedback practices?

1.2 Thesis structure

In the following sections of the first chapter of this thesis, I provide background by describing the educational context of this study, and by defining terms. Understanding the concepts that this thesis is based upon is important for understanding the entire thesis, therefore the theoretical framework is discussed in chapter 2. The theoretical concepts presented in chapter 2 will reappear in chapter 3 through a presentation of previous research. In chapter 4, I give an overview of the methods used in this thesis, including discussions of strengths and limitations of these. The results of the empirical study are presented in chapter 5, and discussed in chapter 6 in light of the theoretical framework. My concluding remarks are in chapter 7.

1.3 The Knowledge Promotion reform-LK06

In order to discuss assessment of writing, it is necessary to present the National Curriculum, as it forms the set of standards that teachers teach and assess by. The current curriculum, the Knowledge Promotion reform (LK06), was implemented in 2006, and signaled a shift from a content-centered curriculum to an outcome-based curriculum. The previous curriculum was content-centered, and limited the influence of local school authorities to secure a common frame of reference for all students, by emphasizing a clear orientation towards content (Dale, Engelsen &

Karseth, 2011, p. 5). In contrast, the current curriculum is aim oriented, and allows a greater degree of freedom to school authorities and teachers in determining content and teaching methods (p. 31). This shift has not been entirely unproblematic, as teachers have found it difficult to translate competence aims into more manageable objectives (Nusche, Earl, Maxwell & Shewbridge, 2011, p. 31). Furthermore, the lack of national assessment criteria results in a lack of shared understanding of what constitutes adequate, good and excellent performances in different subject areas, and can lead to arbitrary and inconsistent assessment practices (Nusche et al, 2011, p. 31). In other words, vague learning objectives and a lack of assessment criteria may impact formative assessment practices negatively, indicating the importance of engaging in a discussion of the possible problems of implementing formative assessment practices.

A new curriculum that is signaled to address some of the issues in the current curriculum is intended to be implemented in 2020. One intention of the new curriculum is to develop the competence aims so that they more clearly describe competence and progression than the aims in the current curriculum. Clarifying the competence aims is thought to make it easier to assess the student's competence, and in this way support assessment practices in schools (Norwegian Ministry of Education and Research[UFD], 2016, p. 60). That is to say, a shared understanding of what excellence, and progression towards excellence looks like, can make it easier to provide good formative assessment, since it means having clearer standards to base the assessment on.

1.3.1. Writing as a basic skill

The framework for basic skills defines five key competencies that are essential to school, work and social life. These skills are *oral skills*, *reading*, *writing*, *digital skills* and *numeracy*, and are considered the foundation for learning in all subjects as well as a prerequisite for the student to demonstrate their competence. This thesis is about the key competency *writing*, defined by the Norwegian Directorate for Education and Training as the ability to express oneself “understandably and appropriately about different topics” in writing (Norwegian Directorate for Education and Training [UDIR], 2012, p. 10). Further, they state that “mastering writing is a prerequisite for lifelong learning and for active and critical participation in civic and social life” (p. 10). The skills required to write comprehensibly and appropriately are the ability to “plan, construct, and revise texts relevant to content, purpose and audience” (p. 10), pointing out that writing is an activity. To develop the ability to plan texts, students need not only use their own judgments, but need feedback from others (p. 10), highlighting the importance of assessment for enabling the students to increase their competencies as writers.

It should be noted that the framework for basic skills was intended to be used for curriculum development, not for teachers. Considering the concern raised by the OECD in 2011 that there was no unified understanding of what constitutes different performance levels (Nusche et al, 2011, p. 31), it is unfortunate that the framework is not used actively, since the framework does contain descriptions of progression. These descriptions could be considered tools for assessment, if they were recognized as such. Considering the focus that has been placed on integrating writing as a key competency in the subject curriculum, and that assessment is a necessary component of learning, it is important to examine formative assessment practices of writing.

While the Knowledge Promotion reform was being implemented, a national writing test in Norwegian was introduced that was intended to be used for identifying strengths and weaknesses in students' writing. In an evaluation of the writing test, severe criticism was partly directed at the implementation of writing as a key competency, and partly at the curriculum (Thygesen, Berge, Evensen, Fasting, 2007, p. 115). The criticism stemmed from an apparent disagreement among teachers of what writing skills students were expected to have at different grade levels. As a result of this disagreement, teachers had highly individualized understandings of what writing and its key components should be. In other words, students could encounter very diverse expectations, depending on their teacher. For instance, when selected texts were assessed by more than one teacher, the results indicated such differences in assessments that the tests could not be considered reliable (Solheim & Matre, 2014, p. 77). The researchers concluded that unless teachers' interpretive community of writing and assessment of writing was strengthened, it would not be possible to execute the curriculum as was intended (Thygesen et. al, 2007, p. 115). While it should be noted that the writing test was implemented in the Norwegian subject, the findings point to the importance of having a shared understanding of what writing entails in the English subject as well, something that is particularly important since writing is integrated as a main subject area.

1.3.2 Writing as a main subject area

The perceived importance of writing is obvious in the National Curriculum; it is both considered a key competency and is a main area in the English subject curriculum. The purpose of learning English is for international communication and cultural awareness, and ultimately to foster democratic values and co-citizenship by promoting respect through interactions between people with different language backgrounds (Norwegian Directorate for Education and Training, 2013). The revised English subject curriculum is divided into four main subject areas with competence aims: *Language learning*, *Oral communication*, *Written communication*, and *Culture, society and*

literature. These should be seen as supplements to each other, rather than separate (Norwegian Directorate for Education and Training, 2013). For instance, competence aims from *Culture, society and literature* can serve to define the content, while *Written communication* can be the vehicle for learning the content, and *Language learning* can be a means of increasing the student's own awareness of the learning process. The importance of writing skills is implicit in all areas of the English subject curriculum, which describes what writing in the English subject entails as follows:

The main subject area includes writing different texts in English in different situations where written communication is necessary to stimulate the joy of writing, to experience greater understanding and to acquire knowledge. This also involves adapting the language to purposeful objectives and to the recipient, i.e. by distinguishing between formal and informal written language. The main subject area involves developing a vocabulary and using orthography, idiomatic structures and grammatical patterns when writing. It also covers creating structure, coherence and concise meaning in texts (Norwegian Directorate for Education and Training, 2013).

This outline of the main subject area of writing indicates that the skills students are intended to develop are relatively complex, higher order skills. Besides being detailed in describing what is expected of writers, the main subject area signals rather high ambitions for students of English in Norway, further emphasizing the importance of studying the role of feedback in the development of writing.

While reading is considered a component of writing in the main subject area called *Written Communication*, it will not be given attention here, since this thesis focuses on exploring challenges of providing formative feedback to students' writing, and students' attitudes towards it. The competence aims from the main subject area *Writing* in VG1 in upper secondary² are ambitious and I understand these aims as serving at least three purposes that can be seen from different perspectives of learning. The first purpose of writing is linked to teaching and learning textual conventions such as spelling, punctuation, vocabulary, grammatical patterns, what is termed *local text levels* in this thesis. These are the aspects of writing that can be seen as "right" or "wrong". Assessment of these can be seen in light of a behaviorist view of learning, and these aspects of text are those that are easiest to correct, since they follow a set of rules that can be consulted.

² Official documents in Norway use 'upper secondary school', rather than 'high school', therefore I do the same.

A second purpose of writing that can be inferred from the main subject area is that of communication, which can be seen in light of sociocultural and social constructivist views of learning. These aims highlight writing to suit the context, such as purpose, text type and situation, including using appropriate language in different situations, such as formal and informal language. These aspects of writing indicate the challenges of writing as well as assessing writing, since contextual factors such as language, cultural and socioeconomic background affect how one expresses oneself.

A third purpose of writing is to acquire topical knowledge and increase understanding of the subject and of writing. This purpose is founded in cognitive views of learning. Examples of cognitive aims are those that indicate that writing is a way to learn and to understand, that writing entails employing strategies, and creating structure, coherence and concise meaning. Additionally, aims that indicate that students need to be able to critically and independently evaluate information, indicate a cognitive view of learning. These aims require being able to identify learning through not only assessing the result, but also monitoring the process, and identifying higher order thinking skills. Consequently, these aims carry with them notions that internal processes can be externalized, and therefore have implications for assessment. In Sadler's (1989) words, making qualitative judgments about qualitative aspects of performance is not entirely unproblematic as it entails a certain measure of appraisal. As these purposes of writing are distinct, they cannot be seen as independent, meaning that the writer usually uses language strategically in a situation, indicating how all purposes are at play. These perspectives on writing indicate that writing – both as a process and as a product – can be a means to demonstrate and measure knowledge, making it all the more important to study assessment practices intended to improve students' writing skills.

1.3.3 The Assessment for Learning project

Following the Knowledge Promotion reform, it was recognized that assessment practices could be improved in schools, and Report no. 16 to the Norwegian Parliament, titled *Early Intervention for Lifelong Learning*, stated in 2007 that Norwegian schools lacked a culture for assessment which had led to students being inadequately followed up, reducing their opportunities for academic development (Norwegian Ministry of Education and Research [UFD], 2007, p. 77). Therefore, a national research project to test 'competence descriptors' for learning objectives was initiated. 'Competence descriptors' are described by The Norwegian Directorate for Education and Training as "descriptions of the quality of what the students master in relation to the competence goals"

(Throndsen, Hopfenbeck, Lie, Dahle, 2009, p. 7). The project, named *Improved Assessment Practices* (2007-2009) found teacher beliefs about assessment to be so diverse that the project recommended the assessment principles in the education legislation should be clarified. For instance, more than half of the teachers felt that effort should count towards the grade, nearly a third of the teachers felt that comparison of students could serve as a basis for assessment, and more than a third of the teachers felt that “a criterion for high goals achievement is that students show great interest for the subject” (Throndsen et al, 2009, p. 13). Following the project recommendations, a national initiative to improve assessment practice was implemented in 2010, named *Assessment for Learning* (Norwegian Directorate for Education and Training, 2015a). In addition, in 2009 the Norwegian Education Act was amended to clarify the assessment principles that provide the basis for assessment in Norwegian schools. The legislation, based on recommended formative assessment practices, states that:

-Students should have a clear idea of what they are supposed to learn, and what is expected of them (Opplæringslova [Education Act], 2009, §3-1).

-Students should receive feedback on the quality of their work or performance, and should be advised on how to improve (§3-11).

-Students should take active part in their own learning process by assessing their own work and progress (§3-12).

The legislation clearly reflects the influence of prominent perspectives of formative assessment, which will be further presented and discussed in chapter 2. In 2015, the Education act was amended further to clarify the relationship between the roles of formative and summative assessment, as this had been an area of dispute in determining end of year grades:

-The competence the student has shown during schooling is part of the basis for the assessment when end of term grades are set. (§3-16).

Still, this relationship between formative and summative assessment remains an issue, and will be further discussed in subsection 1.4.2, as this has implications for assessment practices and how these are perceived by students.

1.4 Defining ‘assessment’ and other terms

The term ‘assessment’ is used throughout this thesis, and must therefore be defined. Bachman defines assessment as “the process of collecting information ... according to procedures that are systematic and substantively grounded” (2004, pp. 6-7). Bachman and Palmer expand on this definition and state that ‘assessment’ is the result or *outcome* [emphasis added] of the process of collecting information, and often takes the shape of a description or a score (2010, p. 20). Sadler uses ‘assessment’ to describe any judgment of a performance or work (1989, p. 120). These three understandings of what assessment is and what it entails, indicate a scale from process to judgement. To further complicate the matter, the terms ‘evaluation’ and ‘assessment’ are occasionally used interchangeably. For instance, certain educational contexts may use the term ‘evaluation’ when specifying certain types of assessment (Sadler, 1989, p. 120). Alternatively, evaluation may be one way of using assessment (Bachman, 2004, p. 9). Whereas ‘assessment’ entails collecting information, ‘evaluation’ “involves making value judgments and decisions on the basis of information” that has been gathered (Bachman & Palmer, 2010, p. 21). This echoes Sadler (1989), and places evaluation on a judgment end of the scale.

Like Sadler (1989), I understand assessment as a judgment of performance, because even if an assessment is used for the purpose of facilitating learning, it involves deciding where the performance is relative to the target performance. With that being said, assessing student texts should entail more than performance judgments. Huot and Perry (2009) problematize the use of the word ‘assessment’, and its strong association with grading and testing (p. 426). Considering the intentions of ‘formative assessment’ to enhance learning, this association is problematic, as will become evident in subsequent chapters.

1.4.1 Formative assessment and Assessment for Learning

Definitions of ‘formative assessment’ have varied in focus and scope. In 1998, Black and Wiliam defined ‘formative assessment’ as: “encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged” (pp. 7-8). A different understanding of ‘formative assessment’ is that of ‘assessment for learning’. In 2002, an expanded definition of ‘assessment for learning’ was provided by the Assessment Reform Group, as “the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there” (Assessment

Reform Group, 2002, pp. 2-3). In this understanding, assessment is intended to promote student's learning (Black et al, 2003, p. 2), and is used by the teacher *and* the learner to decide where the learners are in their learning process, and how they need to work to improve their performance. This understanding of assessment signals a shift from a teacher focus, to a shared responsibility for learning between the teacher and the student.

Based on these definitions, my understanding of the main distinction between 'formative assessment' and 'assessment for learning' is who is responsible for the process of making sure learning occurs. For the sake of clarity, my understanding of 'formative assessment' places this responsibility with the teacher, while 'assessment for learning' includes a student focus on how to best facilitate learning. In this thesis I will use the term 'formative assessment', since the distinction between 'formative assessment' and 'assessment for learning', while being important for theoretical grounding, is not important to the topic of this thesis, which is feedback practices and student attitudes towards these.

Huot and Perry (2009) put forward the idea that rather than using these distinctions, one should be concerned with the goals and roles of assessment. The goal should be helping students improve their writing, and the role of assessment should be to assist students in the process of writing by teaching them to assess their own writing as it improves (p. 426). My understanding of 'formative' assessment is that it has different purposes. First, it can be used to identify particular strengths and weaknesses each student has to adapt teaching. From this it follows that 'formative' assessment can be used to promote learning for the individual student. In the context of this study 'formative' assessment is important for the development of writing skills, emphasizing why quality assessment practice is important. Still, what constitutes quality assessment practice is difficult to define in different subjects, indicating the importance of studying assessment practices, and how these are perceived by students.

1.4.2 Formative Assessment and Summative Assessment

Scriven (1967) is credited with making the distinction between 'formative' and 'summative' assessment, which distinguished between the time of assessment, with 'formative' signifying assessment during a unit, and 'summative' signifying assessment at the end of a unit. This distinction has proved to be problematic. Translated to a writing context, Huot and Perry (2009) describe this distinction as denoting whether the assessment allows the student to improve their text or not (p. 424). The terms are still in use, but 'formative' and 'summative' assessment have

come to represent more than the timing of assessment. For instance, ‘formative’ assessment can be used to describe the process of conducting frequent assessments with the aim of defending a final, and therefore ‘summative’ assessment of what the student has learned (Gardner, 2012, p. 3). This understanding of ‘formative’ assessment resembles the one described in the clarification of the end of year grades in the Education Act, as specified in subsection 1.3.3.

While these two understandings of assessment have different intentions, the first functioning as documentation of achieved competence, and the latter focusing on how assessment can be used to promote learning, the relationship between these two is more complex. The implied dichotomy between ‘formative’ and ‘summative’ assessment ignores the reality that to adapt teaching to meet learning needs, an assessment *of* learning must be made. In other words, formative assessment is about both product and process. Harlen (2012) suggests that rather than seeing ‘formative’ and ‘summative’ assessments as dichotomies, they should be seen on a continuum with ‘informal formative’ assessments on one end and ‘formal summative’ assessments on the other (p. 98). A ‘summative’ assessment can be used to uncover gaps in learning that can be filled by further instruction, thereby giving the assessment a ‘formative’ function. Conversely, assessment termed ‘formative’ may not lead to activities that fill in any gaps, questioning the ‘formative’ effect of the assessment. The continued problematic relationship between ‘formative’ and ‘summative’ assessments can pose problems for teachers and for students, indicating that it is still necessary to discuss this as a possible challenge of ‘formative’ assessment.

1.4.3 Feedback

An important aspect of assessment is providing feedback, which can be defined as “information about how successfully something has been or is being done.” (Sadler, 1989, p. 120). The most influential definition of ‘feedback’ is that of Ramaprasad (1983), who stated that “feedback is information about the gap between the actual level and the reference level of a system parameter which is used to alter that gap in some way” (1983, p. 4). This definition emphasizes three aspects. First, feedback may be focused on any feature (Ramaprasad, 1983, p. 5). Translated to a writing context, this means that feedback may be directed at orthography, ideas, language, structure, or other textual features. Second, for feedback to be useful, it is necessary that there is information on the reference level and the actual level, and third, how to compare the two levels. Ramaprasad states that these three conditions are necessary for feedback in general (1983, p. 5). In an educational context, these three conditions can be seen as the assessment criteria that a performance is measured against, the actual performance, and the method of comparing these two.

Due to the qualitative nature of the reference level, in this case the assessment criteria, measuring the gap between the reference level and the performance is a challenge. Most importantly, only when the information is used to close the gap between the desired outcome and the performance, can it be termed feedback (Sadler, 1989, p. 121). My understanding, and therefore use, of feedback, is that of Ramaprasad's, that feedback should be directed at improving a performance by using a standard as the reference. In this way, feedback can point out areas of improvement, and how to achieve this, indicating how this understanding of feedback fits into an understanding of 'formative' assessment.

As these definitions indicate, the terms assessment – formative and summative – and feedback tend to bring connotations of grading and testing, in addition to more recent ideas of learning. Grabe and Kaplan (1996), point out that writing assessment is often used as both response and the more formal systems of student evaluation, such as grading, and that assessment need not require grading (p. 395). In an attempt to separate these, a distinction is made in Hyland (2003), who separates feedback from assessment. He sees feedback as text response, where the process of writing and rewriting points forward to future writing. According to Hyland, feedback as response helps the writer understand the context, audience and expectations (2003, p. 177). Assessment, on the other hand, is used to measure a performance according to standards (p. 213). In other words, response has an element of process, whereas assessment has an element of product, in Hyland's view. As these definitions indicate, assessment, feedback, and response are important elements of learning-promoting assessment practices. Because I see these terms as elements of formative assessment, I use assessment, feedback and response interchangeably in this thesis. The following chapter will provide the theoretical framework for this chapter, and discuss how formative assessment can be problematic in light of different views of learning, and in light of the demand that formative assessment adheres to the principles of transparency, reliability and validity.

2. Theoretical framework

Formative assessment provides the theoretical framework for this thesis, and the framework that the Norwegian school system wants assessment to be situated within. This chapter addresses the first research question, that asks what formative assessment theory says about ‘best practice’. Therefore, this chapter will present and discuss what formative assessment is and how it relates to teaching and learning. As chapter 3 will illustrate, implementing good formative assessment practices has proven to be a challenge, therefore possible problems of formative assessment will be discussed here. The aim of all education is learning. From a formative assessment perspective, assessment is an integral part of learning. Consequently, since theories of learning often affect teaching and assessment practices, it is necessary for teachers to have knowledge of learning theories, and how these include ideas of what motivates students. Furthermore, as indicated in subsection 1.3.2, theories of learning permeate the curriculum, emphasizing the importance of knowing about, and aligning teaching and assessment practices to these.

An assessment is always ‘of’ something, in this case writing. Since this thesis deals with feedback on written texts, it is important to address what it means to write, since writing is given a great deal of attention in the English subject curriculum, as discussed in subsections 1.3.1 and 1.3.2. This chapter will therefore present theories of learning, motivation, writing, and assessment, because elements of these are included in assessment of writing. Consequently, these theories affect how feedback is perceived by students.

2.1 Principles of Formative Assessment

There seems to be a general agreement as to what good formative assessment looks like, illustrated through commonly accepted principles of formative assessment. While Paul Black and Dylan Wiliam have been influential in promoting formative assessment since 1998, they are vague when it comes to formulating the practical details of formative assessment. Consequently, this section will draw on Sadler (1989) and Hattie and Timperley (2007) for clarifications of what the principles of formative assessment are.

The most influential attempt to design a theory of formative assessment was first undertaken by Sadler (1989), by transferring Ramaprasad’s (1983) definition of feedback to an educational

context. Sadler's article was a response to the problem that feedback on its own did not seem to have a positive effect on the quality of student work (Sadler, 1989, p. 119). Building on Ramprasad (1983), he stated that assessment should guide the student to progress. According to Sadler, feedback is effective when three conditions are fulfilled. First, that the learner understands the goal, which means that the student must "hold a concept of quality roughly similar to that held by the teacher" (p. 120). The second condition for effective feedback is that the student must be able to measure the quality of their work against the standard. The third condition is that the student can take the appropriate action to close the gap between the standard and the quality of their own work (Sadler 1989, pp. 120-121). These three conditions cannot be seen as a linear process, and all conditions for formative assessment must be satisfied simultaneously (p. 121). Most importantly for Sadler is the assumption that assessments made without the participation of the student are insufficient to further their learning. An important distinction between Sadler (1983) and Black and Wiliam (1998, see subsection 1.4.1) is that the former focuses on the feedback loop, whereas the latter focuses on the classroom context. This distinction has implications for theoretical grounding, and ultimately for how feedback is presented by teachers and received by students (see further discussion in section 2.2).

More recently, Hattie and Timperley (2007) sum up these perspectives by stating that for feedback to be effective, it should answer three questions. The first: "Where am I going?" signals that the student should know what the goal of the task is, or what the task should teach them. The second question: "How am I going?" indicates that the student should be informed of how they are performing according to the specified goal. If the teacher does not communicate their idea of a good performance, students cannot be expected to adhere to it. Third, "Where to next?" points to what the student needs to do to make progress (Hattie & Timperley, 2007, p. 86). These questions are termed 'feed up', 'feedback' and 'feed forward'.

As this section has indicated, principles of formative assessment are general, not specific to subjects or the learning of skills such as writing, making the principles a 'one-size-fits-all' solution to assessment. This is problematic when taking into consideration that each subject has its own understanding of what constitutes important knowledge, and how this can be demonstrated, reiterating Bennet's (2011, see section 1.1) call for domain-specific research. It can be argued that providing formative assessment in line with the principles presented in this section is not possible if teachers do not have a shared understanding of what is worth knowing in each subject, and what constitutes excellence and progress. This section began by stating that Black and Wiliam were too

vague in their descriptions of formative assessment. Still, they should be credited with contributing to an increased focus on formative assessment practice, and therefore the reignited calls for establishing a theoretical basis for formative assessment.

2.2 Formative assessment issues

Dysthe (2008) points out that assessment is often discussed independently of the views of learning that teaching practice is based in, indicating a lacking understanding of the close relationship between the two (p. 17). It has long been considered a problem that formative assessment was not established as grounded in theory (Sadler, 1989; Torrance, 1993; Black & Wiliam, 2008; Taras, 2010; Bennett, 2011; Moeed, 2015), to which Sadler's article and subsequent framework was a response (1989, p. 119). This problem is based in interrelated issues, the definitional issue—explained in subsection 1.4.1—and the lack of a theoretical grounding. The definitional issue will be briefly revisited before the lack of theoretical grounding is addressed.

As already discussed in subsection 1.4.1, there are multiple definitions of formative assessment. The introduction of the label 'assessment for learning' has served to increase confusion, as this sends the signal that 'assessment for learning' is something separate from formative assessment. As Taras (2010) notes in reference to various writings by Black and Wiliam: "Dual definitions of formative assessment appear across the literature: one is based on Sadler (1989) and has formative assessment focusing on product assessment...The other is a classroom learning and teaching pedagogy process..." (Taras, 2010, p. 3017). This distinction is important because it denotes two entirely different situations, the first in which a standard is used to judge the quality of a product, and the second involves questions of power and decision-making in the classroom (p. 3018).

The lack of a comprehensive definition of formative assessment raises other issues as well. Torrance (1993) notes that one of the main problems of formative assessment theory is that "there is no single theory of formative assessment" (p. 335), and is somewhat of a circular problem. Without defining and delimiting formative assessment, it is difficult to create a theory to frame formative assessment, and without a theoretical basis, it is difficult to define formative assessment. The reason is that definitions include descriptions of who the main participant is in the assessment loop. If formative assessment is mainly a part of teaching methodology as implied by Black et. al "It has to be within the control of the individual teacher and, for this reason, a change in formative assessment practice is an integral part of a teacher's daily work" (Black et. al, 2003, p. 2). If there

is no grounding theory, then it is left to each teacher to draw on their own views on learning to provide a theoretical basis for the assessment. The lack of one specific grounding theory is a potential problem if teachers are unaware of their views on learning. Furthermore, it is problematic considering that a national initiative towards improving formative assessment practices has been in place in Norway for several years. If the lack of theoretical grounding reduces assessment to a set of individual teacher practices, the intention and effect of the national formative assessment initiative is undermined. In light of this, a discussion of the challenges of formative assessment must necessarily include views of learning.

Various suggestions have been made as to what types of theory can serve to ground formative assessment. For instance, formative assessment could be situated within pedagogical theories of learning. Dysthe (2009) points out that awareness of implicit views of learning is important for the learning environment one wants to develop in schools, and for aligning teaching and assessment to the intentions in the curriculum. Since the topic of this thesis is to discuss possible challenges of formative assessment, different learning perspectives—such as behaviorist, cognitive, and social views of learning—will serve as a basis for discussing how formative assessment can be grounded in theory.

2.2.1 Formative assessment in a behaviorist perspective

Torrance (1993) points out that formative assessment as a concept emerged at a time when education practices were founded in behaviorist theory, and that formative assessment can be seen in light of this. This paradigm is closest to seeing assessment a product-oriented. Greeno, Collins, and Resnick assert that a behaviorist view supports a quantitative perspective of what it means to know and learn, and assessment in this view involves measuring samples of knowledge (1996, p. 37), such as testing factual knowledge that can be memorized but that does not require higher order thinking skills. The concept of ‘mastery learning’ (Bloom, 1968) is an important contribution to teaching, and is based on the assumption that any learning objective can be reached, given enough time and proper instruction. This involves breaking the objectives down into manageable units that contain specific objectives. After the teaching units, the teacher assesses if the objectives have been met. Based on the assessment, the teacher decides if the student can move on to different tasks, or if more time needs to be spent on mastering the objectives (Woolfolk, 2001, p. 221).

This view holds that only observable changes in behavior can indicate learning, and emphasizes external stimuli, such as praise and punishment, as the cause of change (Woolfolk, 2001, p. 201). In an assessment situation, grades can be considered praise or punishment, which is a problem if the student focuses their attention towards improving their grade rather than their performance (Huot & Perry, 2009, p. 427). Glossary tests are an example of a behaviorist view of assessment. The student memorizes words, takes the glossary test, and then moves on to the next unit. While this type of assessment can document student knowledge at a certain point in time, it does not indicate whether the student can use this knowledge later in the proper context. In fact, assessments that focus on checking factual knowledge may encourage surface level learning. It is not until the results of a glossary test are used to identify what the student needs to work more on learning, and which strategies can be used to learn, that the assessment takes on a formative function. As this discussion has indicated, using behaviorist theories for theoretical grounding can be a problem for the learning intentions of formative assessment, since the intention is to achieve more than surface level learning.

2.2.2 Formative assessment and cognitive theories

Newer understandings of formative assessment theory place a great deal of importance on the student in the learning process. Moeed (2015, p. 185) promotes the idea that formative assessment fosters the higher order thinking skills that are fundamental to cognitive theories, such as reflection, understanding, metacognition and expressing ideas. A cognitive view of assessment is concerned with whether students understand general principles and if they employ useful problem-solving strategies (Greeno et.al, 1996, p. 37). This demands that students are able to think about their own progress—termed metacognition—and choose suitable strategies to further their learning. The influence of cognitive theories is seen in §3-12 in the Education Act, which specifies the role of the learner in assessing their own work, and is based on recommendations for good assessment practice. These theories are also present in the National Curriculum through competence aims that specify verbs like “evaluate”. However, the curriculum does not state how these skills should be promoted, and is one of the problems that follows the freedom of an outcomes-based curriculum, versus a content-based curriculum that is centrally controlled. The fact that the curriculum does not state how these skills are to be promoted is particularly problematic for assessment.

Providing feedback on student texts that enables students to develop these skills is demanding, since it entails more than correcting mistakes. Assessment that stimulates cognitive skills must be

directed towards making the students think about how they have solved the task, and what they can do to improve. In other words, the focus must shift from telling the student how to improve, to making the student think about how to improve, indicating how formative assessment may be seen in light of cognitive theories. However, a cognitive view of assessment, demonstrating learning entails being able to demonstrate higher order thinking skills, such as being able to analyze and evaluate information. These skills are neither easy to demonstrate, nor easy to quantify. Assessing cognitive skills means making qualitative judgements of products that are supposed to represent mental processes, which is highly problematic from a formative assessment point of view. What this means is that each teacher decides based on their individual understandings of what demonstrations of cognitive skills look like, what the student should work more on.

2.2.3 Formative assessment and social views of learning

Stobart puts forward that “the learning theory that underpins current AfL positions... is probably best described as social constructivist” (2008, pp. 150-151). This claim makes an important point, though perhaps not the point Stobart intends, namely that ‘assessment for learning’ is closer to having a clear theoretical grounding than ‘formative’ assessment has, according to the distinctions made in sections 1.4.1 and 2.2. Moeed supports this, and emphasizes the student-teacher interaction, and the linkage of new knowledge to previous knowledge as important aspects of formative assessment that are grounded in constructivist and social constructivist views of learning (2015, p. 185). Social constructivism emphasizes that individuals develop subjective meanings of their experiences, and that historical and cultural norms, as well as interactions with others form how the individual sees the world, and how they construct knowledge (Creswell, 2007, p. 21). The distinction between social constructivism and sociocultural theory is blurry at best, and the two terms are sometimes used interchangeably.

The grounding for sociocultural theory came from an opposition towards the focus on the individual as independent of context (Dysthe, 2009, p. 41). Lev Vygotsky (1896-1934), as the foremost proponent of sociocultural theory, sometimes termed a social constructivist, emphasized the interactions between the learner and a more knowledgeable other, such as a parent or teacher (Woolfolk, 2001, pp. 44-45). A central tenet to Vygotsky’s theory is that of the zone of proximal development (ZPD), defined as the distance between what a child can do on its own, and what a child can do with the help of an adult or more able peer (p. 331). Teaching within the zone of proximal development is known as ‘scaffolding’, the basic principle being that the supporting intervention provided by the adult should be in inverse proportion to the competence level of the

student. Social views of learning may be identified in the core curriculum and in the Education act, as they emphasize developing talents with others, and developing social and cultural competence (Norwegian Directorate for Education and Training, 2015b).

Assessment in a sociocultural view of learning stresses the quality of students' participation, and considers assessment practices as basic components that need to be included in all systems of activity (Greeno et. al, 1996, p. 37). Implicit in this view is that the student is not just the object of assessment, but an active participant in different aspects of assessment, such as discussing the aims and criteria and assessing their own work and the work of others. Assessment can be problematic in this view. For instance, there is the issue how to handle the relationship between student participation in the assessment practice and the fact that all assessment must ultimately result in a documentation of achieved competence. In other words, the end of term must result in a grade, a summative assessment. This means that while assessments leading up to the final grade can include peer-, and self-assessment, the final grade is set by the teacher, and ultimately is a documentation of demonstrated learning. The relationship between the teacher and the student is inherently asymmetrical. As long as the teacher is the one assessing the text, the student is writing on the teacher's terms, and the expectations the student thinks that the teacher has. Social views of learning have much to offer for the learning potential claimed for formative assessment, but teachers need to be aware of the power relations between themselves and the students. Furthermore, teachers need to be mindful of the fact that the sum of formative assessments lead to a summative assessment, indicating how problematic the learning intentions of formative assessment are, and why providing theoretical grounding is so difficult.

The most important thing to draw out of this discussion is that different learning paradigms have advantages and disadvantages, and that teachers likely employ principles from all these paradigms. Consequently, it is necessary for teachers to be aware of how they understand assessment in light of how they view learning, as this has implications for how they carry out assessment, and ultimately how students receive the assessment.

2.2.4 Implications of theoretical grounding

As the previous sections have illustrated, formative assessment can be grounded in paradigms of learning, such as behaviorism, cognitive theory, social constructivism and sociocultural theory (Moeed, 2015, p. 185), depending on how one approaches teaching and assessing. This indicates that the time for placing learning in a single camp may be past, and that we must accept that

various types of learning involve processes that can be described in various ways. Sfard advances the notion that rather than a homogenous theory of learning, we must realize that combining different views of learning is necessary because “each has something to offer that the other cannot provide” (1998, p. 10). While there are tensions between these views, it is likely that a theory that borrows elements from others, termed “merged, middle-ground theory... will eventually be accepted as common wisdom and carried into practice” (Shepard, 2000, p. 6). This may be seen as a solution to the problem of grounding formative assessment theoretically, but only if it is made clear how the various views of learning come into play. Translating formative assessment and its theoretical base into practice places a great deal of pressure on teachers, and demands that teachers have a reflexive stance towards their own assessment practices and an awareness of their own embedded views of learning. In other words, as Torrance states, it is necessary to know teachers’ thoughts on the purpose and intention of assessment, and to what extent they are aware of the theoretical assumptions the assessment is built upon and contributes to (1993, pp. 339-340).

Knowledge of views of learning that can serve to frame formative assessment is important for the context of this study for two reasons. The first is that the curriculum incorporates different views of learning, and this is something that teachers need to be aware of, because these ultimately indicate views of what it means to know. Secondly, teachers need to be aware of their own views of learning, since these are implicit in teacher’s responses to texts, and affect how these are received by students.

2.2.5 Transparency, validity and reliability in formative assessment

While the intentions of both ‘formative assessment’ and ‘assessment for learning’ are to increase learning, the word ‘assessment’ tends to be associated with grading and testing (Huot & Perry, 2009, p. 426). One indicator of this association is the notion that the principles of ‘validity’, ‘reliability’ and ‘transparency’ in assessment are seen as basic to all forms of assessment. These concepts were developed for large-scale testing, and it is problematic that some of these principles are better suited for summative assessment practices, yet they are considered to be relevant for all assessment practice. The concepts of ‘validity’ and ‘reliability’ have been problematized for their use in classroom assessment (Brookhart, 2003; Smith, 2003), but these arguments can easily be applied to a discussion of their relevance for formative assessment. In this section, I will present these three principles and discuss if, and how, they are relevant for formative assessment. This discussion is important because, like theoretical grounding, these concepts are in place to ensure

the quality of the assessment, and if they are not applicable to formative assessment, then this is a problem that needs to be addressed.

‘Transparency’ means that an assessment situation must clearly state what is to be learned, and which criteria the performance will be judged by (Nusche et al, 2011, p. 53). The fact that ‘transparency’ is an important part of formative assessment is best demonstrated in that most best practice recommendations point out the importance of stating the goal of the task. Furthermore, the principle of transparency is embedded in the Education Act (see subsection 1.3.3). While the principle of transparency should not be difficult to adhere to, the principles of validity and reliability are more complex, and particularly challenging for formative assessment practices.

‘Validity’ refers to whether what the assessment measures is relevant. This entails that the content in an assessment situation represents the content of the subject domain (Sandvik, et.al, 2012, p. 38). Considering how much autonomy teachers are given in interpreting the curriculum, the demand for validity can be problematic in more than one way. First of all, there is no national syllabus, and teachers can decide what content will be taught and assessed. This means that what is ‘representative for the subject’ is a matter of individual definition, therefore the premise that ‘what is measured is relevant’ must mean what the *teacher* considers is relevant. If teachers are free to decide the content that is to be taught, assess this throughout the year, and the student is chosen to sit the final exam that is created centrally, then the student runs the risk of not having the content knowledge necessary for the final exam, and is at a disadvantage. From this it follows that the teacher autonomy in the national curriculum can be a problem for the relationship between formative and summative assessment. If the argument is that it is not content that is being tested, then one might wonder what *is* being tested. If what is tested is higher order cognitive skills, (see subsection 2.2.2), then there is the question of how to assess this. What this discussion indicates is that the premise that measuring relevant content is a prerequisite for validity, is difficult to fulfill in a formative assessment context where the curriculum is outcome-oriented, and signifies why it is important to continue discussing the challenges of formative assessment.

Another understanding of ‘validity’ refers to the inferences made, how appropriate these are, and how these are used (Nusche et al, 2011, p. 53). Messick (1995) criticizes the premise that validity is the relationship between criteria and results as being simplistic (p. 742). Validity also refers to the chain of interpretations and misunderstandings that occur when competence aims are translated into assessment (Sandvik et.al, 2012, p. 41), indicating how complex validity is, and how its application to formative assessment can be a challenge. After all, if teachers are deciding the

content, as they should according to the curriculum, then the validity of the assessment rests on how they interpret the curriculum. These interpretations include deciding content, then deciding how to demonstrate knowledge of the content. A number of misunderstandings can occur during this process, compromising the chain of validity. The teacher must know what constitutes a 'quality' response to a task. For the teacher to know what constitutes a quality response to a task, they must know which skills the students should be learning, and be able to recognize, describe, and demonstrate a good performance or product. In other words, a high level of subject competence is an essential aspect of assessment competence, and is necessary to increase the validity of an assessment.

'Validity' is also about whether the form of assessment suits the purpose of the assessment, and that the purpose is achieved (Stobart, 2012, p. 233), which is equally problematic. Stobart takes a rather ambitious stance towards the application of validity to formative assessment and states that "if the purpose of formative assessment is to stimulate further learning, then validity is about whether this is achieved" (p. 233). Again, this raises questions. For one thing, one might ask how to measure that further learning has been stimulated. Furthermore, if the intention is learning, this raises issues of how to measure learning. As the subsections 2.2.1 through 2.2.4 have indicated, how learning is measured, or rather *if* learning can be measured, depends on what view one has of learning.

Lincoln and Guba (1985, p. 292) state that a precondition of validity is 'reliability'. This entails ensuring that the assessment is accurate, and not influenced by the person assessing, or the situation (Nusche et al, 2011, p. 53). The goal of making sure that an assessment is reliable is to eliminate variations in results that are due to what is seen as 'irrelevant factors', such as who performed the assessment, which questions are asked, and whether the student is having a bad or good day (Black & Wiliam, 2012, p. 244). These are contextual factors that obviously affect the assessment results in a formative setting. If it is up to individual teachers to decide the content, then the questions they ask to find out if they have learned the content, become very relevant for the reliability of the assessment. Consequently, this understanding of reliability can be problematic for formative assessment.

Stobart finds 'reliability' to be of less central concern in formative assessment (2012, p. 234), though he does not explain why. Harlen (2012, p. 93) dodges this problem by using 'dependability', though not defining what this means. Still, Harlen makes the excellent point that dependability is increased when teachers have a clear understanding of what the goals are and how

to progress towards these (p. 93). In other words, reliability is increased when the teacher has a high level of subject knowledge and a clear idea of what constitutes excellence in a subject, and how it is demonstrated. In any case, two important points can be drawn out of this discussion. The first is that it is essential that teachers have high levels of subject competence, and second, that the most commonly used understanding of reliability is less appropriate for formative assessment practices. This means that perhaps reliability needs to mean something different in formative than summative assessment practices.

One of the guiding principles of ‘reliability’ is consistency in results. This notion is contradictory to the intention of formative assessment, since the whole point is progress. Furthermore, reliability measures differences between students. If all students are taking the same test, then that is possible. But in a formative assessment situation, the intention is to help the individual student progress, and comparison between students is less viable. First of all, each student has individual needs, and learning might not happen in a linear fashion. Secondly, as has been problematized at length, the purpose of formative assessment is learning, and this cannot necessarily be neatly quantified. Smith (2003) suggests a reconceptualization of reliability in classroom assessment. This reconceptualization fits well with a formative understanding of assessment as well, particularly since both have learning as the foremost goal. He suggests that a better understanding of reliability could be ‘sufficiency of information’, which describes whether the teacher has enough information about the student’s achievement to reasonably decide how the achievement measures up to the stated aim (Smith, 2003, p. 30). Smith points out that asking the question of whether there is enough information to make a sound judgement of the performance, brings forth additional questions that the subject domain that is being assessed is properly covered, and whether the task will provide enough information about students’ learning (p. 32). These are questions of validity, indicating that reliability as ‘sufficiency of understanding’ need not ignore concerns of validity.

2.3 Formative assessment and motivation

While learning paradigms are concerned with how and under which circumstances we learn, they also include views on motivation. A simple explanation of motivation is as follows “the study of why people think and behave as they do” (Graham & Weiner, 1996, p. 63), In the context of this study, motivation is important as it relates both to assessment and how assessment is received by the student. When teachers assess, they need to balance their response in such a way that they

point out what the student can do to improve their work, while still providing enough encouragement that the student wishes to continue working to improve. It is therefore important to have an overview of how we are motivated, and how views of motivation are grounded in learning paradigms.

The behaviorist view of learning holds that motivation is seen as derived from outside the learner, and is termed 'extrinsic' (Greeno, Collins & Resnick, 1996, pp. 24-25). In their most basic forms, these may be punishment and praise. School contexts use these types of incentives, in the form of grades, and feedback. The incentive to act is dependent on how the individual responds to the extrinsic motivating factors. For instance, high scores may be motivating, whereas low scores may be demotivating. Rewards are only effective to the point that they answer the internal goals of an individual. Praise is only effective if that is what the student is looking for, and punishment is only effective if the student wants to avoid it (pp. 24-25). This means that positively framed text response may be motivating for some students, whereas critical comments are perceived negatively, though the opposite could also be true. This emphasizes the importance of knowing one's students and how they respond to feedback.

The view that engagement is a result of the relation between the individual and how they organize information is called 'intrinsic motivation', and is grounded in a cognitive view of learning (Greeno et. al, 1996, p. 25). In this view, the person's interest in a subject affects their engagement. Rather than focusing on praise and punishment, a view of motivation as intrinsic is more occupied with finding ways to stimulate the individual's natural tendency to learn. In fact, rewards can have a negative impact on intrinsic motivation, as it may hinder students from taking actions they previously would have taken for intrinsic reasons alone (pp. 25-26). One way of stimulating intrinsic motivation in students when responding to texts, could be to ask the students what they want feedback on. This allows the teacher to provide comments directed at what student want to further develop. However, a possible problem with having the student decide what they want feedback on, is that the teacher may ignore other important aspects that should be given attention in the assessment. In other words, students may not know what they need to improve.

Motivation in a different view of learning is that of 'engaged participation', which is grounded within theories that emphasize social aspects of learning (Greeno et al, 1996, p. 26), such as sociocultural theories. A theory of engaged participation stresses the individual's interpersonal relationships and identity within the contexts they participate in. The contexts a person participates in contribute to forming their identity. If one participates in a community where learning is valued,

the likelihood of being motivated to learn increases. This view holds that we learn from our participation in social practices (p. 26). In assessing writing, the teacher needs to acknowledge the students as writers, and that teachers are participants in their written discourse, rather than merely judges of correctness if they want to stimulate students' motivation. At the same time, students may not know how to improve their texts on their own. This means that while the teacher must acknowledge the student as a writer, the goal is to improve writing, not only increase motivation to improve. As these views of motivation indicate, providing feedback that both encourages motivation and furthers learning is a challenge for teachers, and affects how students receive the feedback.

2.4 Theoretical perspectives on writing

The intention of formative assessment is to help the student gain insight into how to solve the task, as well as become aware of, and learn strategies for problem solving. This means that assessment is part of the learning process, and the teacher is an active participant during this process (Engh, Dobson, Høihilder, 2007, p. 28). Since this thesis concerns formative feedback on student texts, and how this can be used to promote students' learning, theoretical perspectives on writing are important to address. Theories of writing shed light on how the process of writing is intertwined with learning, stressing the importance of having knowledge of writing and learning when assessing written texts.

Cognitive theories (as discussed in subsection 2.2.2), influenced the research on writing that was carried out in the 1970s, and formed the basis of process-oriented writing. The continued influence of cognitive models can be seen in the wording for writing as a key competency in the English subject curriculum: "*Being able to express oneself in writing in English...means planning, formulating and working with texts that communicate and that are well structured and coherent.*" (Norwegian Directorate for Education and Training, 2013). Cognitive theories are problematic in the sense that they describe mental processes that cannot be observed. This means that in assessing texts, it is an oversimplification to assume that the text is a reflection of mental states that can be interpreted by reading the text. Awareness of the shortcomings of cognitive theory in assessment of writing is important for this thesis, as they can explain why teachers might find it difficult to provide formative assessment of cognitive skills, such as reflection and understanding.

2.4.1 The Response Triangle

The 'Response Triangle' (Hoel, 2000) is a cognitive model of writing that describes the process of writing as integrating and organizing components of various kinds and at various levels, and organizes these in a model, see figure 1. This process orientation is well suited for the English subject curriculum, that also describes writing as a process. While this model describes the process of writing, it also describes writing as a product since it divides the text into levels that can be analyzed separately.

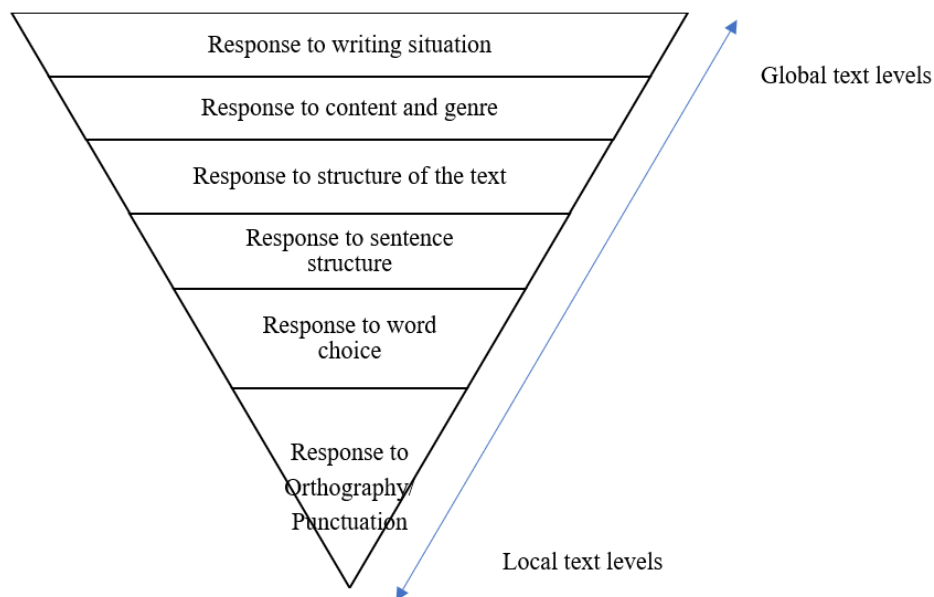


Figure 1 Hoel (2000), after Hillocks (1987)

This model is relevant for this thesis, because it illustrates what text assessment usually takes into consideration, as will become evident in the following chapter. The model illustrates that the lower end of the hierarchy is less cognitively demanding, since these components are more concrete. Word choice and spelling places fewer demands on the writer than global issues such as ideas and audience, which are more abstract. Hoel asserts that when working on lower text levels, it takes a great deal of mental effort to shift over to working with global text issues. In fact, it takes a greater effort than if one were moving from a global text level to a local text level (2000, pp. 33-34). This means that unskilled writers will often struggle with planning and gaining an overview of the global text levels (p. 35). This describes the problem of keeping the overall focus of the text, and of binding the local levels with the global levels. There may be different reasons for this, for

instance, the writer may have too little knowledge of the topic, or there may be too much information to process. Implicit in this model is an understanding that these components of writing follow genre specifications, and these specifications affect how the teacher assesses the quality of the text.

Hoel's model describes L1 writing, but if unskilled L1 writers have difficulties with gaining an overview of the global text levels, it is likely an even more pressing issue for L2 writers. If local text issues are the easiest to work with, and global issues are the most demanding, then feedback should primarily be directed at helping students with understanding how to make sense of global issues. However, in an L2 context local issues may not be that easy to attend to, spelling may be demanding enough if the student only has a weak grasp of the basics of the target language. This emphasizes the importance of adapting feedback to the student while still being aware of theoretical perspectives of writing. Having knowledge about writing means understanding the processes, strategies and knowledge that is important for that specific domain, and is central for effective formative assessment (Bennett, 2011, p. 15). Hoel's model is useful to describe the process of writing, as well as assessing the text, since it provides a vocabulary with which to discuss written feedback provided by teachers, and what the feedback focus is.

It can be argued that the focus on text levels pays too much attention to the formal aspects of writing. This is particularly problematic if teachers do not share an understanding of which formal aspects are most important, because this means that assessment practices will vary according to what individual teachers find important. In addition, this model does not take text types into account, and how different text types and tasks may influence how one works with the text. Furthermore, this model fails to address the learner in the context of writing, the working conditions, the student's motivation for writing, or lack thereof.

2.4.2 The Wheel of Writing

A sociocultural model of writing is 'The Wheel of Writing' (Berge, Evensen, Thygesen, 2016), see figure 2. It is a better tool for the teaching and assessment of writing, because it captures the complexity of writing in different cultural and situational contexts (Berge, et. al, 2016, p. 172). While 'The Response Triangle' can be used to explain writing as a process and a product, the 'Wheel of Writing' does not emphasize the aspect of process as explicitly. The purpose of the model is to create a common understanding of the phenomenon of writing, since this can increase the validity of the assessment.

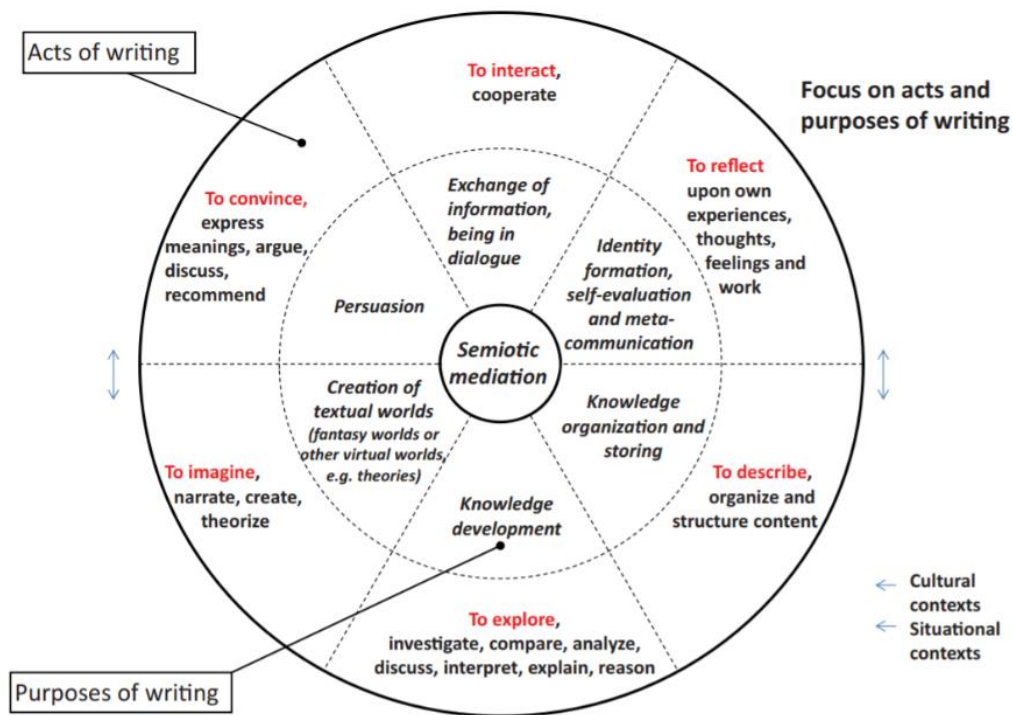


Figure 2 The Wheel of Writing (Berge, Evensen, Thygesen, 2016)

At the center of the circle, ‘semiotic mediation’ refers to how one creates meaning through language (Berge et. al, 2016, p. 175). To do this, the writer uses different acts of writing for different purposes, and these are expressed through the two circles that are beyond the center. The way the model is created, the acts of writing are paired to what is usually seen as the appropriate purpose, though they should not be seen as excluding other acts and purposes. For instance, the act of exploring corresponds to the purpose of knowledge development, but can just as well correspond to the purpose of knowledge organization. The dotted lines signify that the model is not static, and since different acts and purposes can be combined, the text need not be defined within rigid genre understandings of what a text is. The model therefore challenges preconceptions about texts and what writing entails, and can as such be motivating for students (p. 185). At the same time, separating writing from genre can be confusing for students. Some students may need the framework of a genre to structure their text, and focusing on acts and purposes can be too vague for some students.

The second part of the model (figure 3) demonstrates that all meaning making in writing employs tools. This part of the model is important to this thesis because it addresses the resources typically focused on in assessment of writing, namely textual resources like structure and cohesion, and lexico-grammatical resources such as vocabulary and grammar. In this model, it is recognized that these resources follow culture- specific norms (pp. 183-184).

This model is interesting because it focuses neither on genre nor formal features, but whether the text is relevant to its act and purpose, as well as who the recipient is (p. 185). In this respect, the ‘Wheel of Writing’ indicates an understanding of the text as product. This means that assessing the text will be more directed at evaluating whether the students has achieved the purpose of the text, than rigidly focusing on text type and formal features. Since the English Subject Curriculum describes being able to write according to situation and purpose, this model is well suited for that purpose.

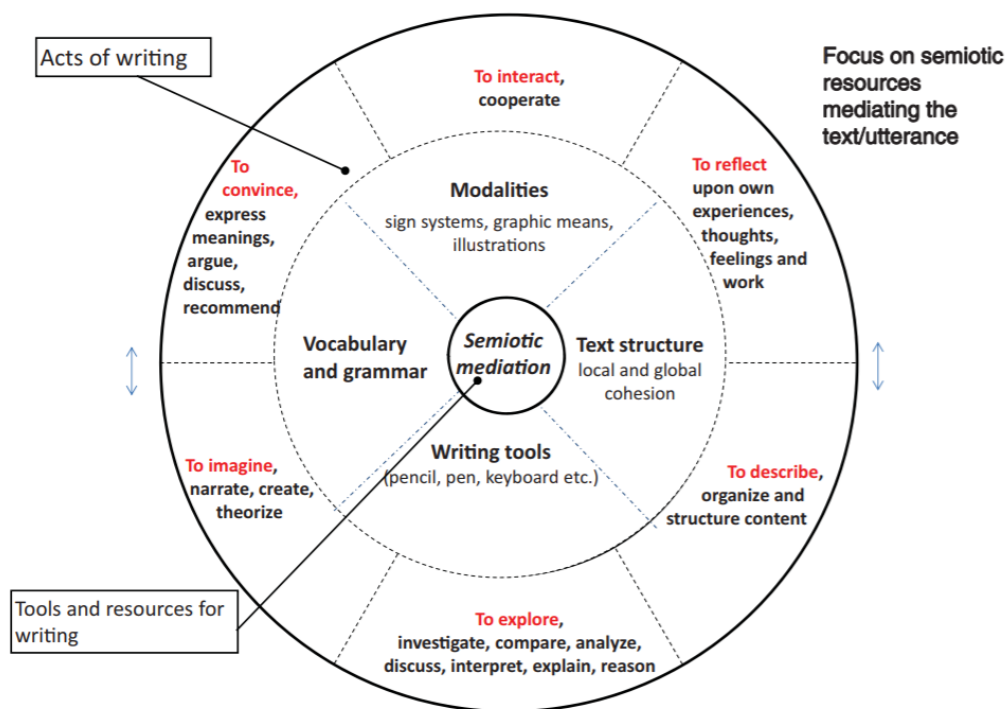


Figure 3 The Wheel of Writing (Berge, Evensen, Thygesen, 2016)

While the English subject curriculum does not specify ‘text types’, this term implies a typology or genre. As long as students may be chosen to sit the central exam at the end of VG1, this means that

those assessing the central exam must have the same flexible view of what a text is. This could potentially raise more pressing issues. For instance, one might ask whether teaching by this model prepares students for the demands of higher education. Also, the model pays little attention to different text levels. While it is positive that the model emphasizes the communicative function of the text, structuring the text is an important part of getting the message across, and the ‘Wheel of Writing’ downplays this aspect. Furthermore, the model is intended to describe writing across subjects, but does not address whether the same model can be used to describe writing in a second language.

What can be drawn from the discussion of these two models is that neither perfectly captures all elements of writing, and that both have their strengths and weaknesses. Both models can be used to describe what writing means—both as a process and as a product—and both are suitable for use with the English Subject Curriculum, as different views of learning can be identified in the curriculum as well as the models. It is therefore important to know about different perspectives of what writing entails, because this has implications for what one focuses on in assessments of texts and how to assess these using generally accepted principles of formative assessment.

2.5 Summary

In this chapter, I have discussed the first research question, which asks what formative assessment theory claims best assessment practices are, and what contextual factors need to be in place for feedback to be effective, such as motivation, transparency, validity and reliability. Since formative assessment is often presented as a set of principles extracted from ‘best practice’ recommendations in scholarly articles, I began by presenting fundamental principles of formative assessment, through the works of Sadler (1989) and Hattie and Timperley (2007). However, as I have tried to make clear, there is a distinction between principles of assessment and a theory of assessment, as the first offers a practical approach, while the latter offers a more abstract view of what formative assessment is and should be. Furthermore, the partly synonymous use of the labels ‘formative assessment’ and ‘assessment for learning’ only serves to confuse rather than clarify what formative assessment means, and would benefit from a clear delineation. Based on the definitions provided, I separate the two by determining the persons involved, so that ‘formative assessment’ is largely a matter of the teachers making teaching decisions based on assessment, while ‘assessment for learning’ includes the student in this process. The confusion that arises from a lack of clearly delineated definitions has been discussed here as problematic since it undermines attempts at

situating ‘formative assessment’ and ‘assessment for learning’ in theory. I have used learning paradigms to discuss theoretical frameworks that formative assessment can be placed within, and how different views of learning carry with them notions of how students are motivated. Finally, theories of writing are presented and discussed, as they offer useful terminology and perspectives with which to discuss why assessment of writing can be difficult. The point of this chapter has been to illustrate how theories of learning permeate not only the curriculum, but also all the practices we as teachers engage in when we assess. Our actions in the classroom are, either implicitly or explicitly, guided by theories of learning, and awareness of these is of the utmost importance. It is therefore necessary that teachers know which theories their assessment is founded in, because their assessment will have a direct influence on students’ learning and motivation, as will become evident in the empirical findings in chapter 5. The usefulness of basing assessment within general principles can be questioned, particularly when they are not positioned within specific subjects. The following chapter will provide insight into how assessment is practiced in various educational contexts, and how students perceive these practices.

3. Previous research

This chapter provides background to the second and third research questions that ask what written feedback teachers provide, and how students perceive written feedback practices. In this way, possible challenges of providing formative feedback may be identified. An abundance of research has been carried out on assessment of writing in the past decades, in various educational settings. Though most studies have been carried out in L1 contexts, there is a growing body of research on writing assessment in L2 contexts as well. The studies presented in this chapter are selected because they address one or both topics of this thesis, namely the types of written feedback students receive on their texts, and students' attitudes towards feedback. In this study, 'feedback types' refers to how written feedback is delivered, with or without grades, and 'feedback focus' refers to whether feedback is aimed at language, termed 'local issues', or at 'global issues' such as structure. 'Student attitudes' refer to their general attitudes towards feedback, how students experience receiving feedback, and how they define this as helpful or not. The research presented here will therefore offer perspectives that range from teacher feedback practice to its reception by students.

3.1 Best practice recommendations

The way written response is framed has consequences for student attitudes towards writing. Response can boost motivation, or it can increase frustration, particularly if the response is not considered useful. The usefulness of response is a matter of subjective opinion, and providing 'one-size-fits-all' guidelines for feedback is a challenge. Dysthe and Hertzberg emphasize that the usefulness of response depends on how response is defined, and what view of learning it is based on (2009, pp. 35-36). Feedback on written texts and how it is received by students is the topic of this thesis, therefore it is relevant to present what best practice recommendations say about responding to student texts.

In two reviews of response literature, Ferris (2014), and Kvithyld and Aasen (2011), sum up advice on how to best provide feedback on writing. First, texts should be provided with response during the writing process (Kvithyld & Aasen, 2011, p. 11), preferably on several drafts, not just on graded papers (Ferris, 2014, p. 8). Drafting can provide the student with a sense of achievement, and can result in increased text competence. Furthermore, drafting can help students monitor their own progress, and is central for becoming independent in the learning process.

Feedback should focus on a range of issues, such as content, structure and language, though not simultaneously. The first drafts should include comments on structure and content. Global issues should be addressed with questions rather than directives, to promote student autonomy (Ferris, 2014, p. 8). Feedback should point out what the student has mastered, to motivate the student, and to provide a model for improving their text (Kvithyld & Eriksen, 2011, p. 14).

Language issues should be addressed in the final draft selectively, rather than comprehensively (Ferris, 2014, p. 8), the distinction being that the first addresses specific errors, while the second points out all language errors. These errors should be indicated rather than corrected, as this is most beneficial to students. Equally important, the errors that are pointed out should be directly linked to learning objectives, and should take the level of the student into consideration, state Kvithyld and Aasen (2011, p. 12). It should be noted though, that there has been a long-standing debate on error correction. John Truscott (1996) triggered the error correction debate by claiming that correcting grammatical errors is ineffective and harmful to students' learning, for both L1 and L2 learners. Truscott claimed that grammar correction is pointless because language structures are gradually learned, and no amount of error correction will lead to a sudden discovery of language structures (2009, p. 342). Since then, much of the research has focused on disproving his claim that error correction is harmful, and Ellis (2009) points out that rather than discuss *if* error correction is effective, current research is concerned with finding out *what types* of error correction are most effective (2009, p. 6). As this section has indicated, what constitutes good response to writing is complex. Since the focus of this study is written feedback that is intended to promote learning and how this is received by students, the following sections will review studies that have taken these two different perspectives on the topic of feedback.

3.2 Written feedback practices

Early studies on written feedback practices focused on how teachers approached texts. In an often-cited study from 1982, Nancy Sommers, studying the commenting styles of 35 teachers from the Universities of New York and Oklahoma, found that teacher comments tended to divert students' attention from their original text purpose, to the teacher's purpose in commenting (1982, p. 149). Students made changes according to the teacher's instruction, rather than those the student considered necessary, this was seen as unfortunate because "students are encouraged to see their writing as a series of parts—words, sentences, paragraphs—and not as a whole discourse" (p. 151). Lil Brannon and C.H. Knoblauch found the same in 1982 when they had 40 teachers respond

to an essay written by a college student and discovered that they assessed the quality of the text from the perspective of an “ideal text” (1982, p. 160). These studies indicate that when teachers assert their authority as experts, students match their writing to what they think the teacher wants. If the purpose of text response is to motivate students to improve their writing and become independent thinkers, students need to have greater authority over their texts. At the time, these studies by Sommers (1982) and Brannon and Knoblauch (1982) signaled a new awareness that teachers need to reflect on their response practices and the theories these practices were based in, issues that are still relevant.

3.2.1 The focus of written feedback

Subsequent studies focus less on teacher control, and more on which textual aspects teachers consider important, through studying what they comment on. A great deal of research on what teacher feedback focuses on has been carried out in L1 contexts. For the sake of comparison of feedback practices, a small selection of Scandinavian studies will be briefly presented here, as they indicate whether the same issues occur in L1 and L2 contexts. Two Norwegian studies indicate that teachers focus mostly on global text levels (Igland, 2008; Bueie, 2016), while one Swedish and one Norwegian study suggest that both global and local issues are addressed in text feedback (Brorsson, 2007; Eriksen, 2017). A Swedish study found that the focus was nearly exclusively on pointing out errors (Kronholm-Cederberg, 2009). As one might expect, considering that these are L1 studies, there is less of a language focus, with the exception of Kronholm-Cederberg (2009). This is confirmed in a large-scale research project from 2012 that found teachers of L1 Norwegian in lower and upper secondary school saw genre and content as more important than language when defining writing competence (Fjørtoft, 2013, p. 99). Considering that an assessment improvement initiative has been in place in Norway since 2010, it is important to study how response is practiced in the English subject, since feedback is considered an essential part of the learning process.

In a study of L2 learners at a New Zealand University, Ken Hyland and Fiona Hyland (2001, pp. 193-194) found that teachers mostly commented on global text levels. However, these were end comments, and a great deal of the in-text comments had a language focus (p. 194). These findings are supported in 2007 by Julie Montgomery and Wendy Baker, who studied L2 learners of English at the Brigham Young University in Utah. They found that teachers provided more feedback on local text issues such as grammar than on global issues, such as ideas, content and organization (2007, pp. 91-93). More interestingly, though, there was a discrepancy between how teachers

thought they assessed and how they assessed (p. 93), indicating the importance and challenges of reflecting over one's own assessment practice.

In a study of two lower secondary schools in Hong Kong, Icy Lee (2008) compared the feedback practices of two teachers from two lower secondary schools in Hong Kong, one high-performance (HP) school, and one low performance (LP) school. Lee (2008) found that both teachers mostly provided error feedback, and that they provided comprehensive feedback in accordance with school policy (p. 149). In a Swedish study from 2015, Karina Pålsson Gröndahl found that the two lower secondary teachers in her study gave roughly equal amounts of feedback on both local and global text levels on texts written in English (2015, pp. 58-65). Except for Gröndahl (2015), these studies show that L2 teachers of English in various countries tend to focus on language correction rather than developing over-arching text issues that can help the student develop their skills as writers. This seems to be the case in Norway as well.

In Norway, where English is taught as an L2, Damir Budimlic (2012) found that the upper secondary teachers in his study mostly commented on local issues, such as orthography and lexical choice (p. 54). The teachers in his study considered local errors to disrupt communication, and thought this type of feedback was the easiest for students to understand (p. 55). In addition, the teachers reported that students and parents expected response at the local text level (p. 55). Budimlic concludes that since teachers are rarely asked about their assessment practices, they find it difficult to articulate the why's and how's of their practices (p. 82), resulting in a mismatch between teachers' beliefs and practices. This raises a few issues, the first being that providing feedback on local text levels to meet the expectations of parents and students is a threat to validity and transparency if the feedback ignores the original purpose of the assessment. The second issue is that if there is a mismatch between teachers' beliefs and practices, then it necessary to address this, since this most likely affects the feedback they provide, and its perceived usefulness.

Three years later, May Horverak's (2015) findings suggested that upper secondary teachers still commented on language mistakes, but seemed to have changed the focus of feedback towards more global levels, such as structure, content and formality. The increased use of global feedback could indicate that an increased focus on formative feedback through the national AfL project (see subsection 1.3.3) has had a positive effect on feedback practices, or it could be a result of the fact that some teachers in Horverak's (2015) study used process-oriented methods. In any case, Horverak concludes that the results of her study indicate that teachers are changing their feedback practices in a more positive direction, since they are using process-oriented approaches (2015, p.

88). The more recent findings of Drita Saliu-Abdulahi, Glenn Ole Hellekjær and Frøydis Hertzberg in 2017, found that the upper secondary teachers in their study tended to provide feedback on local issues in margin comments and global issues in end comments (p. 39). Their findings indicated a clear favoring towards commenting on language issues. Saliu-Abdulahi et. al. suggest that these findings may be the result of either unfamiliarity with formative assessment practices, lack of sufficient subject matter knowledge, or that teachers are hindered by their workloads (p. 49). Whatever the reason, these findings indicate that there is still a need for research on teachers' assessment practices, and how these assessments received by students, in order to engage in a wider discussion about the challenges of formative assessment.

3.2.2 How formative is the feedback?

If feedback is to promote learning, students should have the opportunity to actively use the feedback to improve their performance, as indicated by Ferris (2014) and Kvithyld and Aasen (2011) in section 3.1. Several studies indicate that post-product feedback is most common, and that process-oriented work is least common. Consequently, feedback functions as a summary of a performance rather than indicating potential improvements. The formative potential is therefore lost, as well as transferability to future work. In 2012, Anton Havnes, Kari Smith, Olga Dysthe and Kristine Ludvigsen found that upper secondary teachers in their study seldom gave feedback without grades. Purely formative feedback was rare, because teachers thought students were only interested in grades, while students indicated that they wanted feedback (Havnes et. al, 2012, p. 23). Havnes et.al (2012) conclude that teachers and students need to communicate better.

In 2013, Siv Gamlem and Kari Smith researched students' perceptions of classroom feedback, and found that the lower secondary students in their study perceived feedback as negative if they were not given opportunities for revision. Feedback was perceived as positive if students were given the opportunity to work with the feedback they were given (Gamlem & Smith, 2013, p. 160). These studies indicate the importance of allowing students time to work with feedback, if they are to consider it helpful. Both studies are multidisciplinary, and investigate feedback practice in other core subjects, as well as English. However, recent studies that focus exclusively on the English subject as an L2 in Norway have produced similar results, that teachers tend to provide post-product feedback.

The findings in two case studies conducted in upper secondary school in Western Norway by Linda Beate Vik (2013) and Elisabeth Nyvoll Bø (2014) show that teachers consider process-

oriented work to be a good way to work with texts, but seldom use it with their students. The teachers in Vik's (2013) study agreed that feedback during the writing process was the ideal means of providing feedback, though not all teachers practiced this. Nyvoll Bø (2014) found that most teachers in her study provided post-product feedback with grades, though they too considered process writing to be beneficial. In both studies the teachers cite time constraints as the reason for not employing revisions strategies.

In a recent study, Saliu-Abdulahi, Hellekjær and Hertzberg (2017) found that the upper secondary students in their study were seldom given time to work with the feedback they had received, and teachers rarely had one-to-one discussions with the students on how to use the feedback. Furthermore, the follow-up stage most often included nothing more than students reading the comments, and correcting errors at the sentence level (Saliu-Abdulahi et.al, 2017, p. 42). In most cases, the feedback was considered more a tool for improving the next text, rather than revising the current text. In fact, some of the teachers felt that feedback-assisted text revision could not be fairly graded, and equated this with cheating (2017, p. 43), offering one explanation of why text revision is not used as much.

Horverak (2016) found that the teachers in her study seldom or never have their upper secondary students apply revision strategies. Nearly one third of the respondents in Horverak's study report that assessments rarely lead to text revision, nearly half of the students in her study state that text revision is followed up by subsequent assessments (Horverak, 2016, p. 133). It should be noted that this study is of student perceptions of feedback practices, not observations of practice. Still, since these findings support the findings in previous studies, there seems to be some truth to these claims.

Considering that the teachers in these studies are aware of the benefits of working process-oriented, and that students need time to use feedback for revisions, one might wonder why that teachers mostly still provide post-product feedback. It is therefore interesting to note that in 2012, Lise Sandvik and Trond Buland, while investigating assessment practices in Norway, found that English teachers in Norway conduct thorough assessments of students' writing competence. Furthermore, teachers consider process-oriented writing to be a useful component of good assessment practice. The researchers concluded that teachers exhibit a high level of subject competence in the way they discuss writing instruction (Langseth, 2013, pp. 138-139). At the same time, teachers express uncertainty towards which components of writing they should focus on in assessment of writing, whether it be content, form or structure, and how much importance should

be placed on these elements (p. 139). Teachers also report being preoccupied with fairness in grading, and therefore find exam results to be either a confirmation of, or a mechanism for correcting their own assessment practices (p. 124). These findings are important because they indicate that there is a need for developing a common understanding of what it means to write, in order to improve assessment practices.

The study by Sandvik and Buland also found that VG1 students receive feedback on more extensive assignments, and that these combine grades and comments. However, students did not use these comments in a systematic manner, indicating that assessment practices are not used formatively (Langseth, 2013, pp. 135-136). The idea that teaching and learning are both part of a process is central to formative assessment. If teachers see the value of process, but are not using this in their own practice – for various reasons – then surely this means that there is a need for discussing the problem of implementing formative assessment practices within existing systems of teaching and assessing. Simply put, if everyone agrees on the value of formative assessment practice, it is necessary to discuss why implementing formative assessment in schools is such a challenge.

Based on these studies, it seems that there is room for improvement of feedback practices. These studies have indicated that revision work is not used all that much, and it can be argued that providing feedback is a waste of time if students are not given the opportunity to process the feedback. This is ironic, since teachers seem to find revision work too time-consuming, but not providing revision opportunities makes the time spent giving feedback wasted. Considering the claims for the efficacy of formative assessment, these findings indicate that assessment practices are so varied that they merit further study, particularly since there has been an increased focus on improving assessment practices, which one could assume should have had a positive effect.

From 2012-2016, a large-scale Norwegian study called '*Developing national standards for the assessment of writing – a tool for teaching and learning*' (the NORMS project) set out to improve writing instruction and assessment practices, within a formative assessment perspective. The aims of the study were to define what levels of writing competency could be expected of students after grade levels 4 and 7, and how these defined norms could be of assistance in the teaching and assessing of student texts (Berge & Skar, 2015, p. 9). In the preliminary phase of the study, it was discovered that there were different assessment cultures and practices both within groups and between individuals. There was one commonality, that the informants mostly focused on surface level language issues in their comments, though they were also preoccupied with subject content.

The researchers speculated that teachers found these textual dimensions the easiest to define, since they deal with concrete textual features, and are a way to check content knowledge (Matre & Solheim, 2014, p. 224). Moreover, the teachers tended to make holistic assessments of the texts, and relied on experience when assessing. The researchers assumed this was due to insufficient knowledge of textual levels (p. 226). Teachers also tended to emphasize contextual factors. Consequently, several teachers found it to be difficult and unnatural to assess the text as a text, without taking the student into account (Solheim & Matre, 2014, p. 82), though the researchers do not specify why. This foreshadows a finding in section 5.2. Based on the pre-study the researchers concluded that there was an apparent need for developing teachers' knowledge of texts and metalanguage (Matre & Solheim, 2014, p. 226).

While the NORMS study is from a middle school context, the results are interesting for this study because they indicate the necessity of a shared understanding of what constitutes excellence and progress, and how to assess this, as did the evaluation of the national writing test (see subsection 1.3.1).

3.2.3 Implications of feedback practices

To sum up these findings, section 3.2 through subsection 3.2.3 have shown that in various L1 contexts, teachers comment on both global and local text levels, with some variation. This could be because these teachers are commenting on texts in their first language. The studies in English L2 contexts show a tendency towards focusing on local text levels, rather than global text levels, and that students are generally given little or no time to revise their texts. As will become evident, this has implications for how, and if, students revise their texts. Considering how Norwegian schools have had an increased focus on assessment for learning since 2010, one might have expected different results in the Norwegian context.

Through the studies presented here, a few possible reasons for the varied assessment practices have surfaced. First, the level of subject matter knowledge can have an impact on assessment, particularly if teachers are uncertain of what to focus on in their assessment. If, as explained in subsection 2.2.5, issues of validity are affected by all the choices that lead up to the assessment, then insecurity towards what aspect to focus on when assessing texts, can be a threat to validity. If teachers demonstrate high levels of subject competence when discussing the teaching of writing in English, one might wonder why they are insecure of how to assess writing. After all, teaching writing should be directly linked to assessing writing. In other words, if teachers have a good grasp

of what good writing looks like, and they try to teach this to their students, then they should be able to identify good writing in student texts, in this way strengthening the validity of the assessment. Yet, a large-scale study found English teachers in Norway to exhibit a high level of subject competence. Evensen (2009) points out (see section. 1.1) that each subject has its own ‘culture of knowledge’, meaning that there is a shared understanding of what is worth knowing, and how to operationalize and demonstrate this knowledge. Teachers need to have a clear—and preferably shared—understanding of the aims, and what progression towards these looks like (Harlen 2012, p. 93). Since the content of an assessment must represent the content of the subject domain, as mentioned in subsection 2.2.5, uncertainty of what to focus on can have implications for the validity of the assessment.

Second, factors outside the assessment situation, such as exams, school policy, workload, and expectations from students and parents may have an influence on how teachers assess, and what they focus on in their assessments. If this is the case, then these are serious issues, because they imply that teachers are assessing based on factors that are separate from the assessment situation and criteria. The finding that teachers are using exam results to confirm or correct their own assessment practice supports the notion that there is not a shared understanding of what constitutes quality in the English subject. Furthermore, using exam results to confirm or correct one’s own assessment practice points to the problematic relationship between formative and summative assessment purposes. Considering these findings, that external factors are influencing how assessment is carried out, it is noteworthy that not one of the studies mentioned here problematizes what this means for the validity or reliability of the assessments. Avoiding a discussion of these issues ignores the importance of securing the quality of assessment practices.

Third, teachers may not always be aware of their beliefs, including what they see as important to assess. A possible result is that how teachers assess and how they think they assess may not align. Considering the implications of assessment—on motivation and learning—for students receiving the feedback, there is a need for more research on teacher feedback practices, and how they are received by students.

3.3 Student attitudes towards written feedback

Preferences towards teacher’s comments can arguably be said to be highly subjective. Still, as recipients of feedback, student attitudes towards feedback should not only be investigated, but

should also be used in comparison with assumptions of what constitutes good feedback practice. It follows that if students find the feedback to be less than helpful, or difficult to understand, then time spent providing feedback is time wasted for the teacher and the students who receive the feedback. Student attitudes towards feedback in English is an area that has been paid little attention in Norwegian assessment research, therefore studies from L1 contexts will be briefly presented, before addressing studies from L2 contexts, including Norway.

In an English L1 context, Richard Straub found, in 1997, that college students prefer comments that are specific and elaborate, that point out problems and indicate ways to improve their writing. Furthermore, they preferred positively framed comments, in the form of guidance, though not controlling. Similarly, Melanie Weaver found in 2006 that business college students considered feedback to be unhelpful if it was too general, lacked guidance, focused on negative aspects, or was unrelated to assessment criteria. In a Norwegian L1 context, Agnete Bueie found in 2016 that students in lower secondary school preferred detailed comments presented as advice or guidance. However, in contrast to the study by Straub (1997), Bueie's respondents found comments that were critical of their ideas helpful, and preferred comments on a global text level (Bueie, 2016, p. 14). L2 studies seem to focus more on general comment categories, not unlike the categories in Hoel's (2000) response triangle, explained in subsection 2.4.1.

In an English L2 context, Lee (2008) found that students wanted more written comments on future texts. The high performing (HP) students wanted more feedback on content and language, while low performing (LP) students favored comments on organization and language. Roughly 20 percent of the HP students wanted more error feedback, while the same was true for nearly 30 percent of the LP students, though about 30 percent of the LP students wanted less error feedback. Furthermore, HP students were more positive towards the feedback, found it more useful, easier to understand, and easier to correct than LP students. In sum, HP students were more positive towards the feedback they received than the LP students. Still, few students wanted to revise their texts. Lee suggests that assessment practices at both schools are to blame, since they encourage students to be passive by employing single draft assignments, and little use of student participation in the assessment process (2008, pp. 156-157).

Until 2016, research on student perspectives on assessment in the English subject in Norway had not been investigated much, but has since then been a popular area of study. As one of the first to explore the topic of assessment seen from the student perspective in the English subject in Norway, Tony Burner (2015a) found that students in one lower secondary school perceived

comments to be focused on local text issues, resulting in students' increased focus on local text issues during revision processes. In addition, the students found feedback to be vague and negative, which had a negative effect on their motivation to revise (p. 634), and that they wanted more positive feedback (p. 633). Like Burner (2015a), Maria Vågen (2017) found that lower secondary students in her study tended to revise comments that were directed at local text levels, and that students were positive towards assessment practice as a tool during the learning process, but only if the feedback was detailed and exact (p. 69). Furthermore, the study indicated a need for more explicit feed forward.

In a recent study of upper secondary students in VG1 general studies in Norway, Saliu-Abdulahi (2017) found that students like receiving both general feedback and specific feedback, though there is a slight preference for specific feedback. In this study, 'general feedback' is defined as feedback that "sums up the quality of writing", and that suggests how the text can be improved, focusing mainly on content and structure (Saliu-Abdulahi, 2017, p. 137). This definition corresponds to what others here have termed the global text level. 'Specific feedback' is considered feedback that refers to language issues, like grammar, vocabulary and spelling (p. 138), or local text levels. Students report preferring specific feedback because these issues are easier and less time consuming to address, which is important for two reasons. The first is that, according to the students, most are not given sufficient time to work with the feedback they have received, since it is mostly provided post-product, and second, that language seems to be given a priority in grading. In other words, bad grammar might reduce their grade, whereas structure is not given as much emphasis (pp. 141-143). Furthermore, students felt that the comments alone were not sufficient, and preferred when they were combined with one to one discussions with the teacher to clarify. Since several of the studies here have found that students often receive feedback with grades, it is worth mentioning findings from two studies on the reception of feedback in combination with grades.

3.3.1 Grades and feedback

Many studies find that feedback is often provided with grades, therefore it is worth looking into what previous research has found on student attitudes towards grades and feedback combined. A 2009 study by Anastasyia Lipnevich and Jeffrey Smith of US university students found that the students in their study saw the usefulness of grades from two perspectives. If the goal was to complete a course spending minimal amount of effort, grades were helpful, because they indicated how much work to put into the course. However, if the goal was to learn, then the students

recognized grades as getting in the way (p. 356). All students found detailed feedback to be the most effective form of feedback (p. 364). This study indicates the importance of motivation in the learning process, and confirms the idea of intrinsic motivation presented in section 2.3. In addition to motivation, the student's performance level seems to have an impact on the student's perception of grades. In a Norwegian lower secondary school context, Burner (2015b) found that the students in his study differed in their perceptions of grades, depending on their performance level. The low performing students preferred grades on all texts, as did some of the average students. However, the high performing students were more positive towards adapting to feedback practices that were more in line with formative assessment, and which downplayed grades (p. 63). Academic maturity as well as motivation and proficiency level could be the reason, though Burner hypothesizes that students were positively affected by the teachers. In other words, any changes in student attitudes were impacted through the change in teacher practice, rather than through conscious effort to change student perceptions of the usefulness of formative assessment practices (2015b, p. 70). This finding suggests that while it is important to improve teacher assessment practices, it is equally important to train students in the use of formative assessment to further their learning. One may also wonder if students are accustomed to assessment practices that render them passive recipients of feedback, placing a great deal of responsibility on students to motivate themselves, and further emphasizing the importance of studying how assessment practices are perceived by students.

From the studies presented on student attitudes towards written feedback, one can conclude that comments that are specific and detailed are the most helpful comments, both for students of English as an L1 and as an L2. Furthermore, many students want more feedback. How the comments are framed is important, since comments that are perceived as negative seem to have a negative impact on motivation. When discussing text levels, it seems that particularly L2 students prefer comments directed at local text levels, as these are the easiest to attend to. Regarding grades, there is a tendency that motivation and performance play parts in students' attitudes towards these, so that students that are motivated are less preoccupied with grades than those that are struggling or that are less ambitious. At the very least, these findings point to the necessity of discussing the challenges teachers face in formative assessment practices, since these have such an impact on student's perceptions.

3.4 Summary

This chapter began by indicating the perceived importance of writing in the English subject curriculum, and advice on how to best provide feedback that promotes improved writing skills. The studies presented here indicate that assessment practices are quite varied, both in L1 and in L2 contexts, and in various age contexts. The most recent studies find that teachers tend to provide feedback at both local and global levels. When contrasted with previous studies, this shows a tendency towards an increased balance between commenting on local and global issues. It still seems to be an issue that students are not given sufficient time to work with the feedback they receive, and that feedback is mainly given post-product, even though teachers seem to be aware of the advantages of revision processes for improving writing. A lack of revision time has implications for how students perceive the feedback they receive, and seems to result in students not utilizing the feedback they receive, making the time spent on providing feedback somewhat wasted.

From the studies presented here on student perceptions towards feedback, one can surmise that for feedback to be considered helpful by students, it needs to be clear and specific as opposed to vague, which means that students should know how to act on the feedback. Furthermore, the feedback should be positively framed, as negative feedback is demotivating, and it needs to indicate ways that students can improve their work. However, critical feedback is not always seen as negative, so this can be a matter of perception. In other words, a critical comment may be perceived as negative, but not necessarily. It is also worth pointing out that students correct what the teacher points out, so if most comments are directed at local text levels, that is what the students will focus on. Finally, students need to be allowed time to work on revisions, but whether they do so depends on their motivation to perform in that subject, and their performance level.

Since the topic of student attitudes towards feedback is still relatively understudied in Norway, this thesis will contribute to expanding our understanding of what feedback teachers give, and how this is received by students, to enable a wider discussion of why formative assessment is challenging to practice. This makes it all the more important to study *how* feedback is framed, and *what* is focused on in feedback. Furthermore, it is important to examine if the feedback is compliant with principles of assessment that are considered to promote learning, and which views on learning are implicit in assessment practices. The next chapter will present and discuss the choice of methods employed in the process of collecting data.

4. Method and Materials

Considering the importance of providing quality feedback to help students improve their writing skills, and the fact that previous research in Norway shows assessment practices to be quite varied, there is a need for more research on how formative feedback is provided and how students receive this feedback. To this end, this study explores what types of written feedback students in grade VG1 in the general studies program receive on texts written in English, and how students perceive this feedback. To do this, I have analyzed and categorized written feedback provided by teachers on a sample text, and put together a focus group to identify student perceptions of and attitudes towards written feedback using the written feedback provided by the participating teachers. The informants in this study are four students from the general studies program from one upper secondary school in Hedmark County in VG1, and seven teachers that are either currently teaching at grade level VG1 or have done so within the last two years. Five of the teachers are from the same school, and four teachers have master's degrees in English. The student group is made up of two male and two female participants.

In this study I took a qualitative approach to explore what types of feedback upper secondary students receive on written drafts in the English subject, which types of feedback students find the most useful, and why students find certain types of feedback useful. Qualitative research is preoccupied with *why*, and seeks to gain a deeper understanding of the perspectives and attitudes of the participants (Patton, 1990, p. 13). I chose this approach because it allows me to study the issue of feedback in depth. Typically, data collection in qualitative research consists of three different types: interviews, observation and written documents (p. 10). The research design in this study is a sequential design, meaning that it is made up of different sequences that are deliberately staged, and build upon the findings in the previous stage. These stages are described in more detail below (see also figure 4).

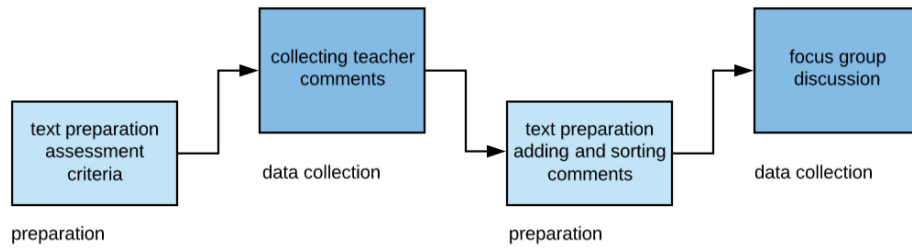


Figure 4 Data collection process

4.1 Phase 1-Teacher comments

Data collection was carried out in two sequences. The methods used in the first sequence of the study were modeled on the methods employed in the studies by Straub (1997) and Bueie (2016). This phase intends to answer the second research question, what types of written feedback teachers give on student texts, and why. To find out what kinds of written feedback teachers give students on texts, I constructed a response situation, in which 6 of the 7 participating teachers were asked to provide written feedback to an authentic text written by a student in grade level VG1. The text was written as a first draft response to a process writing task, with assessment criteria that the student was familiar with, and was obtained with informed written consent from the student who wrote it, and anonymized to ensure their privacy. The teachers were instructed to give the same type of feedback that they would with their own students. Included with the student text is the task it aimed to answer, and criteria for assessing the text, as well as instructions for the teachers, and a request that the teachers include a note on their thoughts on why they chose that way of providing feedback, and, if relevant, what types of problems they encountered during the process. After the teachers in this study returned the texts with feedback, the comments were analyzed and categorized according to types of feedback. Out of this analysis, I selected comments to create a new version of the student text using the margin comments function in Word.

I chose this method because collecting feedback from several teachers on the same text allows for comparison of what types of feedback teachers give, since the text functions as a common reference point. The student text that formed the basis of this sequence was chosen because it represents a mid-level performance, and has possible areas of improvement on both local issues, such as concord errors, and global issues such as structure. This allows me to see what kinds of areas teachers deem most important when providing feedback, which is important for the overarching aim, to find out how to find out what types of feedback students are given, and what kind of feedback students find useful. The feedback was treated as a source of primary data, which was

analyzed and categorized and selected to ensure a broad range of comment types within the categories of local and global comments. This meant revisiting the comments from all feedback samples I collected, and identifying which aspects were most commented on, and selecting comments that were good representations of their categories.

4.1.1 Analysis of teacher comments

While other categories appeared as the comments were analyzed, the teacher comments were first sorted according to the categories employed by Straub (1997) and Bueie (2016). Bueie (2016, p. 48) problematizes the process of categorization. In her dissertation, Bueie (2016, p. 48) intended the categories in her analysis of teacher comments to correspond to the study by Straub (1997), but found that all comments to some degree relate to all categories used by Straub, which were: *focus*, *specificity* and *mode*. *Focus* indicates whether the feedback is directed towards global text issues or local text issues. *Specificity* refers to how specific a comment is, and Bueie (2016) rates these on a scale of *vague*, *medium* and *specific*. *Mode* as a feedback category signals how the comment is presented. This is where Bueie's study slightly diverts from Straub's categories. While Straub operates with the categories *criticism*, *praise*, *imperative*, *advisory*, and *open and closed questions*, Bueie adds *instruction* as a mode of feedback. Figure 5 illustrates some of the categories I used in my analysis of the teacher comments.

Mode	Imperative	Comment that tells writer what to do
	Advice	Comment that suggests a course of action
	Question	Comment worded as a question. May be open, this allows writer to choose course of action, or closed, which asks for specific information
	Correction	Comment that provides the correct form or answer
	Instruction	Comment that shows the writer how to correct an error
	Explanatory	Comment that explains <u>why</u> something is done well
	Praise	Comment provides a positive response

Figure 5 Category examples

My final categories were based on Bueie (2016), but while analyzing the comments, two comment types appeared that did not fit into any of the categories by Bueie. I termed these: *correction* and *explanatory*. *Correction* is a comment type that provides the student with the correct form of a grammar issue, or the correct answer, while *explanatory* comments provide students with the reason for a positive response. In other words, the latter explains to the student *why* something is well done, thus providing something more than praise. Categorizing the comments provide a vocabulary with which to discuss comment types and student preferences towards these. Figure 6 indicates how the comments were categorized. See appendices 5 and 6 for complete set of categorized comments.

2nd paragraph				
C14 Could you use a linking strategy to start the paragraph with?	C14	Global	Medium	Open question/advice
C15 There is one reason, but there are many reasons 😊	C15	Local	Vague/Medium	Instruction/correction
C16 at	C16	Local	Specific	Correction
C17 Rephrase this sentence	C17	Local	Vague	Imperative
C18 though	C18	Local	Specific	Correction
C19 I am/ you are /he she it is/ we are /you are /they are 😊	C19	Local	Medium/specific	Instruction/correction
C20 You present both the number of native speakers and the presence of English online as reasons for why English is a lingua franca. I agree! You do state that there are many reasons, but you only list two. Can you think of and refer to more reasons? Also, good use of sentence connectors, like 'the first reason is'!	C20	Global/Local	Medium/specific	Explanatory/Advice/ Open question
C21 person	C21	Local	Specific	Correction
C22 Try to cover other uses of English than just the internet.	C22	Global	Specific	Advice

Figure 6 Categorized comments-sample

4.1.2 Strengths and limitations of using teacher comments

There are aspects of this method that may influence the results, as well as the validity and reliability of this study. These will be discussed in the following section. The first set of possible limitations addresses the first sequence of data collection: the constructed response situation, the sample text, and the teachers assessing it. In trying to find out what types of written comments teachers provide on texts, I constructed an artificial response situation by using an authentic student text. The text, with assessment criteria and instructions, was piloted by three teachers who were not participating in the study. The pilot test revealed that the original assessment criteria were too detailed, since they were divided into levels in a rubric. Therefore, the assessment criteria were simplified and divided into *content*, *structure* and *language*, based on the criteria Bueie (2016) used, but modified to better suit competence aims in VG1 in the English subject curriculum. It can be argued that altering the assessment criteria corrupts the data, since the criteria may have had an

impact on how the student constructed their text. However, since an important aspect of this thesis is how students respond to written comments, it is important that the teachers that were providing written comments found the criteria to be helpful rather than a hindrance.

By using a text from a student that the teacher is not acquainted with, I created an inauthentic situation. Most teachers who know their students will likely assess a text with individual student issues in mind, thus providing an assessment tailored to the student and their needs, this may also include using codes that they expect the students to be familiar with. Moreover, the assessment situation dictates how teachers assess. In other words, feedback on a text that is not process-oriented, and that will not be revised, is likely to look different from feedback that the student is intended to use as part of a process. Furthermore, teachers may feel self-conscious about knowing that their comments will be read by many, and that may affect how they assess the text.

I could have used authentic student texts with the feedback they have received as a way to collect and categorize teacher feedback. This would circumvent the issue of teacher bias, particularly if they were written before this study began, because then I would be sure that the assessments were representative of how they usually practiced assessment. Those might have been more difficult to obtain, and would have added the issue of securing permission to use the accompanying texts from students, as these would be necessary for understanding the context of the comments. Considering that text type and writing situation affects how the teachers provide feedback, which in turn affects focus of feedback, comments from various authentic texts would not be comparable to each other. Since the teacher comments gathered in this study were directed at the same text, the data should be comparable. Furthermore, since the teachers all provided feedback to the same text, there may be a measure of consistency that increases the reliability of the study.

Reliability refers to the replicability of a study (Cohen, Manion & Morrison, 2011, p. 199), and to the production of consistent results across multiple occasions of use (Denscombe, 2014, p. 271). The method of data collection is clearly described, and so replicating the study should not be difficult. As to the question of consistency, it addresses whether the same results can be produced at different times given that the other factors remain equal (p. 71). The provision of teacher feedback is an individual practice, and it is unlikely that new teachers commenting on the same text would comment in the exact same manner. Increasing the reliability of a replicate study would therefore necessitate using the same student text with the teacher comments I have inserted.

The purpose of asking this number of teachers to provide feedback to the student text was to ensure a wide range of comment types to show students, as part of the purpose of this study is to find out what types of feedback students prefer. While sorting and analyzing the comments, I found that the feedback provided by the teachers was not as varied in focus as anticipated, and yielded few comments at the global text level. To resolve this issue, it was necessary to find one more teacher to provide feedback to the text. This is further addressed in chapter 5. Receiving feedback from a teacher at a different school provided me with the opportunity to compare feedback practices to those within the same school.

The student text with the selected teacher comments formed the basis for parts of the focus group conversation, in which I investigated student attitudes towards written feedback, and which made up the second sequence of this study.

4.2 Phase 2- Focus group interview

In the second sequence of this study, the aim was to find out students' attitudes towards written teacher comments to answer the third research question. The focus group interview method was chosen to elicit student responses towards the teacher comments. For the focus group conversation, I used an interview guide that was divided into four main areas, and that was partly based on Saliu-Abdulahi (2017, p. 152), and inspired by Lee (2008, p. 164), see appendix 7. The first questions were warm-up questions concerning the students' own attitudes and thoughts on writing in English. I included these as a way to get the conversation started. The second set of questions addressed the types of feedback the students typically encountered. These questions invited the students to describe the feedback practice they were familiar with, as well as what they do with the feedback. The third section of the interview guide was less structured. To find out what types of feedback they found useful, the students were first shown the 7 different samples of teacher feedback where they evaluated how they preferred feedback to be presented. In other words, the students were invited to discuss if they preferred hand-written comments, digital margin comments, or in-text comments. The students were then given time to read the task, assessment criteria and student text, after which they were shown the student text with the feedback comments that had been selected from the feedback samples. I read each feedback point out loud and asked the students to share their thoughts on whether or not the comment was understandable and useful. The final set of questions were concerned with other factors concerning feedback, such as timing, and motivation.

I chose this method because it is a good way to explore attitudes (Denscombe, 2014, p. 188), which is the purpose of this sequence. In this project I planned for 4 participants, because I wanted to ensure that all participants were able to contribute to the discussion, and there is no consensus on what the perfect number of participants is, though literature suggests between 4 and 12 participants (Halkier, 2010; Cohen et.al, 2011; Denscombe, 2014). Furthermore, I wanted to ensure that there were enough participants to ensure a range of opinions, but not so many participants as to be unmanageable (Denscombe, 2014, p. 188). Moreover, 4 participants are few enough that the respondents should feel comfortable speaking in a group.

Since the number of participants in this sequence of the study is so low, it cannot be based on random selection (Halkier, 2010, p. 30). Following my request, the participating teachers chose the focus groups students, since they knew the students, and could make informed decisions that could affect the group dynamic positively. Using existing social groups is increasingly common (p. 34). Some advantages of this can be that drawing on existing social networks allows participants to elaborate on the perspectives of other participants, since they may have similar experiences. At the same time, it is important to be aware that social power structures may have an impact on what the students say (p. 35). Furthermore, since the data produced by focus groups is dependent on the interactions between participants, the groups should “neither be too homogenous, nor too heterogenous, because this may cause conflicts or suppress understandings” (p. 30, own translation). Having the teachers select students with varying levels of academic achievement and backgrounds ensured that students with different levels of proficiency participated. I considered this necessary to ensure a variety of opinions, to expand my understanding of how students perceive written feedback.

4.2.1 Analysis of the focus group material

The data was analyzed by using thematic analysis as described in Braun and Clarke (2006), because it is a useful tool to distinguish and report patterns within a data set (p. 79). I used this method of analysis because the focus group interview contained somewhat broad questions, and because I was largely asking the students to speak freely on a topic. This meant that the conversation did not follow a linear fashion, and thematic analysis allowed me to find the broader issues of what the students discussed. Using thematic analysis entailed working with data from the focus group interview over several phases. In the first phase, I transcribed the focus group interview, and familiarized myself with the material, looking for response patterns, and keywords that were most common. I chose keywords based on the patterns I identified, and then searched in

the document for instances of those keywords, and color coded these. In the second phase, I made a list of ideas, and put these into mind maps, before making codes that could encompass several ideas. Codes identify a feature that can be assessed in relation to the phenomena that is being studied. This process was demanding, because the focus group discussion tended to divert from the questions I was asking. After the data had been coded initially, I combined and sorted these into broader categories. I then reviewed the themes, to decide whether there was enough data to support the themes, if new themes occurred, or if themes had to be collapsed into fewer themes. Finally, I defined and named the themes.

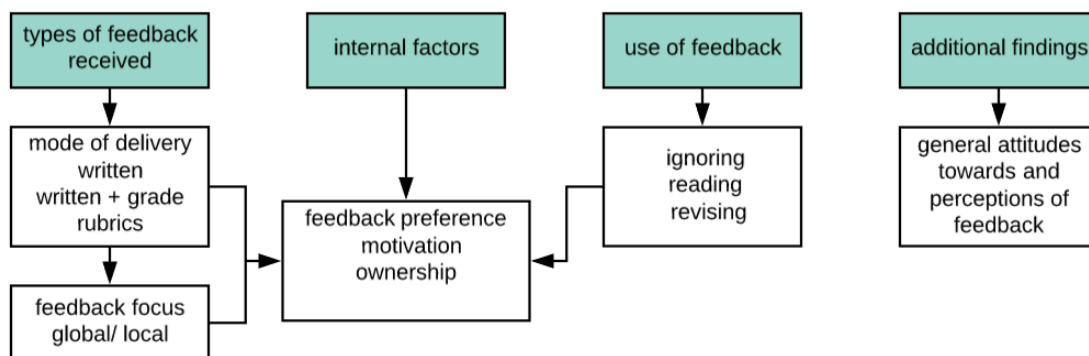


Figure 7 Focus group categories

The focus group analysis resulted in four main categories, see figure 7. The first main category was (i) *types of feedback received*, which had the subcategories (1) mode of delivery, and (2) feedback focus. The second main category was (ii) *use of feedback*, with one subcategory detailing how students act on the feedback. Between these two main categories are feedback preferences and motivation which are subcategories to the main category (iii) *internal factors*. The final main category was (iiii) *additional findings*. This category contains findings that concern the focus group students' general attitudes towards assessment and how they perceive feedback practices at their school.

4.2.2 Strengths and limitations of the focus group method

I chose to carry this study out at grade VG1 in the general studies program in upper secondary school because there are 10 schools in Hedmark county that offer this program, and this could potentially ensure many participants. However, soliciting a response from schools proved to be

challenging, as most schools failed to reply, despite repeated attempts on my part. Of the ten schools that were invited to participate, one school showed an interest in participating.

The fact that only one school was able to participate in this study prompted a change of method. Initially, I had planned on administering a survey to all students at grade level VG1, asking them to rate the usefulness of teacher comments. This would have yielded large amounts of quantitative data that could have shown a tendency of what types of comments students showed a preference for. However, surveys are better suited for larger scale studies, and one school does not fill this criterion, as there are only three classes in the general studies program. To remedy this, I could have included vocational studies, but they have the amount of teaching hours spread out over two years, which means that their subject progression is not identical to that of general studies. The study would then be a comparison of two student groups, which would have altered the scope and focus of this thesis. Furthermore, a survey would not have provided the reasoning for student preferences, and might as such be less useful on its own. Therefore, I decided to run a focus group, as that would not only provide me with the student perspective on feedback, but it would allow me to find out *why* certain types of feedback are considered useful or not. Furthermore, the dialogic nature of a focus group could provide me with information I may not have sought to find initially.

I could have carried out individual interviews, but found that a focus group interview could provide a greater amount of data within a short period of data collection than individual interviews. Furthermore, the focus group can elicit a greater number of expressive and emotional points of view than individual interviews (Kvale & Brinkmann, 2009, p. 162). I therefore expected the focus group to provide both the individual responses of the students in this study, and how they understood the feedback, both individually and collectively.

Since the focus group is dependent on the participants and their willingness to communicate, selecting participants was important. For this reason, I requested that the teachers selected students that communicated well with each other. At the same time, I did not want the focus group to be too homogenous. Acting on these suggestions, the teacher selected 2 boys and 2 girls, all with different levels of achievement in English.

I held the focus group interview at the beginning of January 2018, so that the students would have some experience with receiving and working with written feedback. The interview was held at the schools the students attend, for their convenience, and so that they would be in a familiar setting. I wanted the students to answer in depth, and I wanted to minimize the chance that they would

refrain from answering if they were unsure of how to express their thoughts in English. Therefore, the students were given the choice between speaking English or Norwegian, whichever they were most comfortable with, and chose to speak Norwegian.

There is always an element of bias with qualitative data like a focus group interview. First of all, the questions I ask, and how they are phrased, can limit the student responses. As I mentioned in my framework for focus group analysis (subsection 4.2.1), the students did not always stay on topic. While I took care in formulating the original interview questions, there is the possibility that spontaneous follow-up questions could have been leading. Furthermore, I should have had a tighter structure of the focus group interview. I allowed the students to speak freely, so that those who wished to answer could do so. This resulted in two of the students dominating the conversation. I tried to solve this by prompting the two other students to answer, though they sometimes declined to do so.

As recipients of assessment, students usually have the opportunity to clarify misunderstandings. In this case, the students have no relationship with the ‘imaginary’ teacher, and as such, must rely on their own interpretations of the comments. The fact that the text is not written by the students in the study means that they are not likely to experience ownership of the text, which may affect how they interpret the comments. Furthermore, it is important to be aware that feedback preferences are highly subjective, and as such a study like this can only show possible tendencies.

The validity of a study like this depends on if the research instrument measures what it intends to measure, which in this sequence of the study is student attitudes towards written teacher feedback. To claim validity in research, there are several principles that need to be met. For instance, the data should be descriptive, the principal source of data should be gathered from its natural setting, and the data should be analyzed inductively rather than using predetermined categories (Cohen et.al, 2011, p. 180). I argue that the data presented in chapter 5 is descriptive, and particularly the focus group material is gathered in its natural setting, which is the school context. The focus group discussion analysis is analyzed inductively, allowing categories to appear from the material rather than using predetermined categories. However, validity also concerns whether the results can be generalized to a wider population (p. 186). It cannot be claimed that the results of a small study like this can be generalizable, but they may support previous research.

4.3 Ethical Considerations

This study, and the methods employed, have been approved by the Norwegian Center for Research Data (NSD), see appendix 1. The teacher distributed consent forms with information on the project in December 2017, and obtained signed consent from the students and their parents in advance, which were sent to me as pdf files by e-mail. The students were reminded before the focus group began that participation was entirely voluntary, and that they could revoke their consent at any time.

The focus group interviews were conducted in, and transcribed in Norwegian. The sections of interview that are quoted in the chapter on findings have been translated. However, translating word for word obscured the meaning in some cases. Therefore, I considered it necessary to paraphrase in places. Furthermore, the quotes that are included in the findings chapter are made more concise, because the transcript is of verbal conversation, and contained many discourse markers and fillers such as ‘like’, ‘in a way’, and ‘sort of’. These not only made the transcript tiresome to read, but in many cases the discourse markers were so frequent that it made it difficult to understand what the student was communicating. These two changes, paraphrasing and removing discourse markers, while making the quotes easier to read and understand, may make the statements seem stronger than they were intended by the students. Finally, it must be acknowledged that I have interpreted the findings, and these interpretations will be affected by my preconceptions on the topic of feedback.

5. Findings

The overarching aim of this study is to discuss possible challenges of providing feedback in line with formative assessment theory. To this end, theoretical perspectives of formative assessment were presented and discussed in chapter 2 as a response to the first research question. Previous research on assessment practices and student perceptions of these were presented in chapter 3 to provide background for the second and third research questions. This chapter will answer the second and third research questions by first presenting which types of written feedback students receive on their written texts. And second, by presenting student perceptions of the usefulness of feedback, and what factors influence these perceptions. In the first section of this chapter, I will therefore begin by presenting findings from the first phase of data collection, the teacher comments. In the second section of this chapter, I will present results from the focus group interview to establish what types of feedback students found most useful, and what their³ attitudes towards feedback are.

5.1 Teacher comments

In this section, I present findings on teacher comments, and answer the question of what types of written feedback teachers give on texts, and their explanations for these. The teachers in this study had very different ways of commenting, including the focus of feedback, the commenting format, and the degree to which the comments could feed into future writing. All of the teachers from the one school named contextual issues as problems that influenced how they assessed the text. The following sections will describe these findings in more detail.

5.1.1 Feedback focus

The first finding of relevance in this study is the feedback focus. The 7 teachers in this study provided the student text with a total of 293 comments. Most of these comments were directed at the local text level, such as spelling or grammar. Out of 293 comments, 250 comments were directed at the local text level, these make up roughly 85 percent of the total amount of comments. Comments directed at global text levels, such as structure, content and ideas, made up just below one tenth of the comments. As mentioned in chapter 4, the study included 6 teachers at the

³ In this chapter, I try to avoid using gendered pronouns for the purpose of anonymity.

beginning of the project. Since most of their comments were directed at the local level, I asked an additional teacher from a different school to provide feedback to the text. This teacher is number 7, and as table 1 shows, this teacher commented more at the global level than at the local level.

Teacher	1	2	3	4	5	6	7
Local text level	64	72	27	42	37	4	4
Global text level	4	6	1	3	2	5	8
Margin comments	2	5	28	0	39	0	12
In-text comments/markings	0	0	0	3	0	0	0
End comments	7	2	3	1	4	9	1
Total number of comments	73	79	31	46	43	9	12
Word/Pdf	Pdf	Pdf	Word	Word	Word	Word	Word

Table 1 Overview of teacher comments

5.1.2 In-text comments

The teachers had somewhat differing ways of providing feedback. Teachers 1 and 2 used interactive pdf files, as shown in table 1. The text was given handwritten feedback using a stylus pen on a digital copy of the text, and had a visual appearance that separated these two from the rest. Both corrected comprehensively, that is, they corrected or marked all errors in the text. Teacher 1 tended to provide explicit comments, such as the correct form or word where there were errors, whereas teacher 2 used metalinguistic feedback, such as codes signifying what needed to be attended to. For instance, G signaled a grammar issue, MW means missing word, and WC means word choice, as illustrated in figure 8. The teachers that used interactive pdfs tended to provide a higher number of comments in total; teachers 1 and 2 had nearly twice as many feedback points as the other teachers. Furthermore, these two teachers provided more than twice as many comments on the local levels than the other teachers. In fact, as figure 8 indicates, they tended to either correct or somehow highlight every local level mistake.

Twitter, snapchat and Instagram. Distance is not a problem to communicate. People can communicate even they (is) thousands of kilometers away from each other. English is a useful language when people wants a conversation with another guy who have different language on internet. For example, in handle between countries or chat with a foreigner.

History, the local culture and language in British colonies became weaker, because of "immigrants" from the Great Britain, and the British government. Native people got forced to learn British culture and English. For example, children to native people who lived in Australia was taken from their parents by the government. They was sent to a place to learn English and British culture. After 2-3 generation, many native culture and language disappeared, and the English became the major

Figure 8 teacher 2

Teachers 3, 5 and 7 used margin comments in Word. These teachers had distinct ways of commenting, though they all used the comment function in word. For instance, teacher 3, shown in figure 9, used codes marked yellow or red, yellow signaling that something had been added, and red signaling that something needed to be corrected.

kilometers. Around 1/4 of the total world population lived in the British empire, and used English as an official language to communicate. However, the Empire doesn't exist anymore. But the language and the culture the British empire had left behind, can still be seen in every land where the empire had ruled.

In the end of 17th centuries the British navy successfully defeated his enemies, and became a strong power over the sea. The empire grown fast because of that, and became to the largest colony empire in the history. Colonization helped the spread of English to the world. The industry revolution gave the British empire technology, and goods from colonies became important for production in the Great Britain. Handle between colonies and Great Britain became more than before. Many British people

05. December 2017
A different term would be better here.

Svar Les

See comment above

To grow – grew – has grown ▾

colonial

industrial

This is «norwegian» - ▾

Figure 9 teacher 3

Furthermore, teacher 3 had suggestions for improvements, such as grammatical rules and word choice, whereas teacher 5 had empty margin comments with words and phrases were marked, but no explanatory text. Teacher 5 was the only teacher who explicitly used the assessment criteria, and marked in yellow the issues that should be followed up on. Teacher 7 marked entire paragraphs, and commented on both local and global issues within each paragraph. Teacher 4 had a marking system that included marking words or chunks that contained errors in yellow, and writing in-text comments in Word, like in figure 10.

In the end of 17th centuries the British navy successfully defeated his enemies, and became a strong power over the sea. The empire grown fast because of that, and became to the largest colony empire in the history. Colonization helped the spread of English to the world. The industry revolution gave the British empire technology, and goods from colonies became important for production in the Great Britain. Handle between colonies and Great Britain became more than before. Many British people moved from the Great Britain to colonies to discover a new life. Those immigrants brought the English culture and language from the-Great Britain to colonies, and spread it all around the world.

Can you think of a transition sentence that "introduces" the following paragraph?

Figure 10 teacher 4

Teacher 6 did not make corrections or mark errors, but collected all feedback on a separate paper. These findings from teacher 1 to 5 indicate that the feedback practices at this school are quite varied, and subject to personal preference rather than a unified assessment practice.

5.1.3 The formative potential of end comments

All teachers provided end comments, ranging from a few sentences to several paragraphs. The end comments tend to sum up the quality of the work, and in some cases point to areas of improvement. The findings here illustrate that the end comments are quite varied in focus and length, ranging from nearly a full page to three sentences at the end of the student text, see figures 11 and 14.

Language positive:

You use a varied vocabulary and several formal terms to express your thoughts.

Your sentence structure is varied, you have started sentences with many different words and terms in addition to using several linking words, such as «But on the other hand» and «the other reason»

Your English is idiomatic. You are fluent. You make mistakes, but are very understandable and your mistakes are not great enough to make you difficult to follow or understand.

Language to work on:

Some of your ideas are not precise enough because you need to explain them more. For example «kids getting English lessons in school» why is this there? You don't explain your point.

Explain why kids learning English in a non-English speaking country is important before going to your next point.

You need to work on plural and singular forms of have/has and is/are.

We say «the USA, but don't say «the Great Britain», however we do say «theUK» it is a bit tricky to remember

Work on spelling. Especially when using the same term again and again, such as empire

Figure 11 teacher 1

In chapter 2, principles of formative assessment are presented. One of the principles is that the student should be informed of how they are performing according to the goal (see section 2.1). This can be defined as the degree to which teachers refer to the assessment criteria. The findings indicate that the assessment criteria are used to variable degrees. For instance, figure 11 demonstrates that teacher 1 actively used the assessment criteria. The same main areas are defined: 'language', 'content' and 'structure', which reflect the assessment criteria that are included with the text. This teacher has further divided each main criteria section into *positive* and *needs more work*. Teacher 3 had a similar style of providing end comments, though more condensed. Like teachers 1 and 3, teachers 2 and 6 provided end comments in a similar style, but without specifying clearly which criteria sections they were referring to, these were implicit in the terminology. For instance, one specified criterion was: "You answer the questions raised in the task". Teacher 6 refers to this in the following way:

- You answer important aspects of the question by pointing at some reasons to how English has become the most important language in the world, and you also point at some ways in which English is used as a lingua franca.

Figure 12 teacher 6

Teachers 4 and 5 did not refer to the assessment criteria in their end comments. In fact, teacher 4 did not refer to the assessment criteria in any way, neither in the text nor in the end comments, as illustrated in figure 14.

Check the highlighted words/phrases for language. All of those have some sort of a small mistake (focus on word choice, concord, verb tense, word class (verb, noun, adjective and adverb). Try to fix as many as possible on your own to begin with and ask if you need help!

Figure 13 teacher 4

Teacher 5 marked the assessment criteria that the student needed to work more on in yellow, using the sheet that was included with the task:

Structure:

- Your text has a suitable title
- Your text has an introduction that presents the topic
- The main body of your text addresses the topic in a logical fashion
- Your text has a conclusion that sums up the main points from your text

Figure 14 teacher 5

5.2 Teacher reflections on assessing

How teachers assess is influenced by different factors, such as contextual factors. In the instructions included with the student text, I asked that the teachers provide comments as if the imaginary student was working process-oriented. Only after I had sent this to the teachers was I informed that they did not work process-oriented, because it was too time-consuming.

I asked the teachers to write a note on problems they ran into while providing feedback. Unsurprisingly, a few felt it was difficult to respond to a text without more context, such as who this student was, and which language background they came from, what the student had learned about writing, and what their problem areas were. As teacher 2 stated:

I found it difficult to assess something from the outside like this. It's one thing to do it in a summative fashion, but here the focus was formative assessment, and for that a lot of information is missing compared to what I am comfortable with.

This finding is not surprising, considering that this was an inauthentic situation that has removed contextual clues, which emphasizes the importance of contextual factors in assessment. Teacher 4 dealt with this problem by envisioning a situation where the assessment was based on the assumption that the students spend a fair amount of time writing at school after a first round of feedback, as if they were working process-oriented. Teacher 4 states that

I try to highlight parts of the text where something is a little off, and let the student try to correct as much as possible on their own, while still giving them the opportunity to ask if there is something they do not understand.

Teacher 5 had a similar strategy, and stated “I usually mark in the text, the student looks at it and makes changes, and then I have a chat with each individual student at school once they have had time to look at their texts.”

Teacher 3 changed strategies upon seeing how much there was to address:

If this were process-oriented, I would address one content-related aspect, one structure-related, and one language-related aspect instead of leaving the whole thing red! I try to reflect that in the end comments. But it is difficult when you have not been in dialogue with the student or know the context...

Teacher 6 found it to be an interesting task, and stated that “I obviously need to learn to limit the feedback I give my students. Pick my battles so to speak. This means that there is a lot I have not pointed out – like insufficient citing of sources, for instance”. These findings indicate that contextual factors, like knowing the student, and the amount of time devoted to the task, are important factors for how teachers assess student texts. Furthermore, these findings indicate the challenges teachers face in deciding how much to comment on, and where to direct their attention.

5.3 Focus group results

The focus group interview was divided into questions that concerned students' attitudes towards and experiences with feedback in general, and on how students perceived the comments in the example text. The focus group conversations revealed that, while students did have some preferences towards certain comment types, they were generally pleased with most comments, this

was due to the assessment practices they were accustomed to. Furthermore, they wanted more feedback on their own texts, and found having meta-conversations about feedback helpful.

5.3.1 Student preferences of teacher comments

When shown how the teachers provided feedback on the student text, the students showed a preference for receiving feedback in Word documents with the comments in the margins, and with end comments that sum up the quality of the work. They found the pdf versions to be too disorganized. However, they quite liked the fact that the pdf versions were comprehensively corrected, as they very much wanted to receive feedback to improve their texts. Still, they stated that the comprehensively assessed texts were overwhelming at first sight, and that they would need to get used to this. They also liked the idea of using color codes for feedback on local issues. When discussing end comments, students agreed that longer end comments that specified what was done well, and what to work more on, were preferred.

The students were generally positive towards nearly all comments, both at local and global levels, with a few exceptions. The least favorable comments were mostly from the categories *imperative* or *correction* directed at local text levels such as the word choice items shown in figure 14:

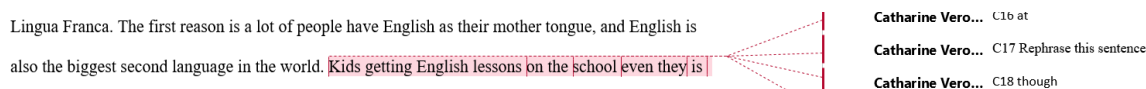


Figure 15

Comments like these were considered unhelpful because they were difficult to understand. For instance, comment 4 (C4) is a one-word comment like C16. The students found this comment to be unhelpful because it does not specify what to do, as this exchange indicates:

STUDENT 4: 'behind', I didn't quite understand that.

STUDENT 3: I didn't understand that either.

STUDENT 4: I mean, the teacher has to give a proper explanation of what they are asking for.

STUDENT 2: Wait, is it that you are supposed to insert that word, to replace that word? Because it goes down there, doesn't it?

STUDENT 4: Maybe it's left, left behind maybe?

STUDENT 2: But then they could have written..

STUDENT 1: Elaborate?

STUDENT 2: Add behind or something like that?

INTERVIEWER: So this is a little unclear?

STUDENT 1: Yes

STUDENT 2: Because here it looks like you are supposed to exchange ‘left’ with ‘behind’, in my opinion, since the whole word is marked.

However, not all comments on local issues such as word choice were viewed negatively. The comments on word choice that were seen as positive were those from the category termed ‘advice’ that explained why the word was used incorrectly and that supplied one or more correct alternatives.

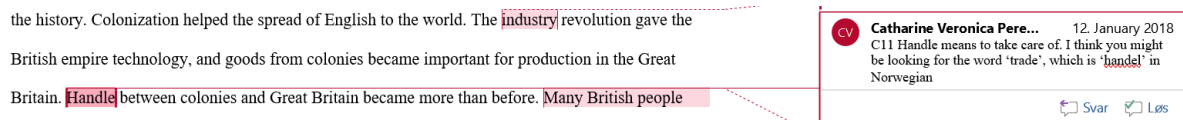


Figure 16

That’s nice, because it gives you an explanation, and you sort of, like sometimes if you’re for instance in an exam or whole-day test, then you write a word, and then you press that word, and you might get synonyms, and then you just choose one, that sounds, well, that sounds nice, but you don’t know what it means, or you look it up in the dictionary because you don’t know the word for it, and so it’s nice that it’s explained, so that you know what she means.

As these two exchanges indicate, comments directed at word choice need to be rather explicit if they are to be considered helpful. However, comments directed at grammar issues, were considered most helpful if they included the grammatical rule rather than the correct answer. This way, they can practice the rules of grammar, but still have to think on their own. In other words, not all comments at local levels need to be as explicit.

At the global levels, students want feedback that specifies what needs more work and why, with examples and explanations. They want feedback that is specific without being controlling, and that makes them think, rather than giving up the answer.

I like that they ask questions, because I think that even though it’s nice to be told what to write, you have to think on your own, because it is a lot about being able to form your own text with your English language, if one can put it like that.

Their preferred comment type is comments framed as questions, these are perceived as helpful and positive, and praise is seen as motivating when it is specific. In other words, they want to know *what* they have done well, not just that something is done well. When comparing their comments on C12 and C13, both categorized as directed at global levels and medium specificity, it becomes clear that the wording has an impact on how the students perceive the comment.

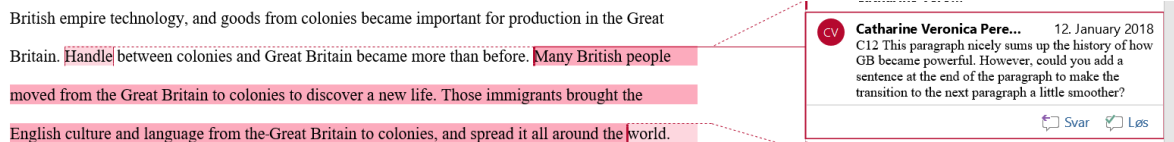


Figure 17

STUDENT 1: I like that one too.

STUDENT 4: It's in a way those that are filled out, that have an explanation, for instance examples, but also a question you can work with.

STUDENT 2: And then you get ... it's okay for a middle aim achievement perhaps, but if you are, if you want to improve, and are aiming at getting higher up, then it's a good way to...

STUDENT 4: That it's a possibility. And that you are given the opportunity, to do that, and I think that maybe the teacher takes into account, if someone is at a grade 2, and has many mistakes, they might not include the things that are necessary to achieve a top grade, they take into account your level, and include hints to how you can improve.

Note that the final comment indicates that students share the notion that context is important for the teacher when assessing. This shared notion could be because teachers at some point have said that they take individual needs into consideration.

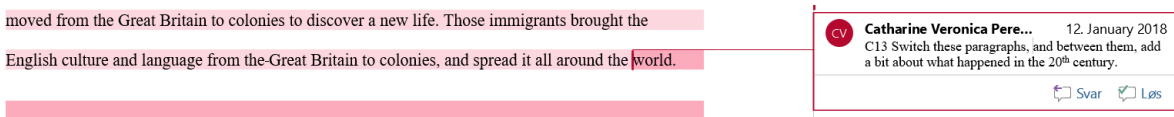


Figure 18

STUDENT 4: There the teacher wants you to put that in. I kind of think that is good, but if the task doesn't ask for this directly, I think it is up to the student to choose what fits into the text, but it could be nice to have something about that, but I would have worded it differently if I were the teacher. It sounds very strict.

STUDENT 2: It's not a very long text, so it might in a way be so that maybe the text, it is the teacher's way of saying that if the text had been a little longer you could have added this, and the text would be even better.

These findings indicate the importance of providing feedback that is specific enough that students can understand what they are supposed to do with the feedback, if the feedback is to be used for future reference. Furthermore, the teacher must balance their feedback, so as not to provide too many answers, but rather ask questions, as these give the student something to work with.

5.3.2 Student perceptions of feedback practices

Parts of the focus group interview centered on attitudes towards written feedback, and the students' experiences as recipients of feedback. The students said that assignments are often given pass or fail, with no indication of how they have performed, or using a rubric with achievement descriptors. If the assessment was provided with rubrics, the students would receive a rubric with descriptions of aim achievements ranging from low to high aim achievement, where the teacher would mark where the student's performance is according to the aim descriptors. These types of feedback made students insecure of how they had performed, and were difficult to use.

High, middle, low, sort of. And we don't usually get comments then? We only get comments if we get grades.

Yes, and if it is an assessment that only says approved /not approved, then you might have received a grade 2, then you have no idea where you are, but even more specifically, if you have a low/middle/high aim achievement, and you receive a middle aim achievement, for instance, then you might think you are at a 4 +, if they add plus or minus, but then you might be at a 3-, because it is still within that scale, and that's quite a difference, almost 2 and almost 5, so it's difficult to find out where you are, and I think that makes it difficult to work with.

No, we usually get [feedback] without grades, so we have just gotten 'approved/ not approved' and 'completed'.

Based on these quotes, it seems that this is a typical mode of receiving feedback. Written comments are rare unless they are accompanied by a grade, and all students expressed discontent with any other type of assessment than grades in combination with written comments. Still, they felt that the written feedback they received did not always correspond to the grade they received. When asked if the feedback indicates how their performance is, they state:

I think that's why we feel unsure, because that's what we notice, that when we receive comments, they don't always correspond to the grade, so that's why we're not sure.

Sometimes you can get a comment like, yes, look at that, but otherwise everything is good, and then you get a four plus or a five or something like that.

Well, the first thing I look at is the grade, right? And then I read the feedback, and then it might say, I can find it here...yes, for instance, here it says that I follow the instructions so that the text has a nice

structure and good job. That I use linking words so that the text is well connected, and then it says a that my word choice is a little off now and then, but that the language is mostly good as well, too (...) and then on the rubric nearly everything has been checked off, except for one thing, meaning that I have done this well. And then I got a five on this. I think I should have been given a five plus, since she has only marked that I need to change one word.

The grades were seen as ‘hard evidence’ of their level of proficiency, and the students worried they would have the wrong impression of their level without a grade. Furthermore, the students stated that grades provided motivation to keep working, though there were different reasons for this. One reason was that grades let the student know where they are, and that provides motivation. In other words, a lack of grades provides no incentive to improve.

Because I think it makes us unsure of where we’re at. Because then we’re sort of sitting around, waiting for our term grade, and that has no, I don’t know, motivation to work, since that grade means a lot if you get a three, then that’s what you get, then there’s a reason you got that three, and you want to improve.

Furthermore, a poor grade on an assignment provided motivation to improve their performance. This was particularly true for the male participants who had a more pragmatic attitude towards grades and using feedback than the female participants. For instance, unexpectedly poor grades were motivating for the male participants, but if they were satisfied with the grade, then they were not motivated to improve their performance, and therefore did not read the feedback.

If I get a good grade, yes. Because then I think that there’s not much to improve, right?

Sometimes, for instance, if I have other subjects, like science, and I know I haven’t studied or understood anything in class, then I know what I can improve, so there’s no point in reading the feedback.

While they preferred grades with feedback, the degree to which they used the feedback varied. Two of the students reported not always reading the feedback, while two of the students always read the feedback, though always after looking at the grade first. If the students were happy with the grade, they might not read the feedback, but if they were unhappy with the grade, they would read the feedback to find out how to improve their grade next time. Furthermore, they all stated that they would read the feedback if they received only feedback.

I don't look that much at the feedback. I do read them, but I can't remember right now. But the first time we wrote an English text in upper secondary, I asked the teacher how to improve my text, and she said that there wasn't much to improve, right, and I thought that then it was a 5, or higher, but then it ended up lower, but the thing is that you don't know **what** to do better, when you don't know what to do better **on**, so that is what's difficult when writing texts.

Like, if had only received written comments, that might give a different impression of how we (...) were doing.

Two important factors for reading and using the feedback were that they were not given time to read the feedback, and that they were rarely given time to work with the feedback in class.

Well, it depends, sometimes I don't, not always, but I might not have time to read the feedback because we might we get the grade or feedback during school hours because we have class. But I may have time to read when I get home.

We don't do much with the feedback, it's not like we sit down and change what is wrong, we could have been given some time at school to do that.

Yes, and maybe been given some time make corrections as well. We do take it with us, if one reads the feedback, and it states what one has to do better.

Furthermore, the students recognize the value of working with feedback.

I feel like we often write a text, and it is about a topic or a sequence, or lesson or whatever, then you work with it, hand in the text, and then start a new topic. And then I feel like they don't make the time to think about what one should do better. Working with the text afterwards too, I think that is actually kind of important for learning.

Like in other subjects, we are told that we need to do repetition, that that's why we have homework, that we need to repeat what we are doing. So, I think that if we get the text back and do things, we kind of get to repeat it, and we learn better.

The students were all positive towards feedback, and expressed awareness of the learning potential feedback provides. In fact, all students felt they received too little feedback, which is interesting, considering that they reported using the feedback only occasionally. When asked if there were other things that could help the students become better writers, they felt that talking about feedback had been very useful, and that improving the quality of feedback of the feedback they received could help them improve their writing. Their ideal method of

receiving feedback is, they say, feedback that tells them what they have done well, and concrete feedback on they can work more on, and conferencing with the teacher.

From this short session, it has actually helped, because I have looked at other texts, and so I know what I can do later to improve (...)

Feedback that contains what I can work on, what I have done well, and points out, very concretely what I need to work on in my text, because that's what's kind of sloppy, that it's kind of general in a way, right?

And then sitting with your own text with the teacher, and like, sitting and looking through, and pointing, that you get proper feedback.

These comments point to an important issue, which can be something of a dilemma to deal with, namely the relationship between grades and feedback. First of all, the students want more feedback, but they also want grades. Second, if the students receive grades, and are satisfied, then they might not even read the feedback. It seems then, that the best solution is to provide feedback without grades, though this makes the students feel insecure about how they are performing. This indicates just how important the quality of feedback is for the student to understand where they are in their learning, and for the student to want to continue working to improve their performance.

5.4 Summary and synthesis

The findings presented in this chapter are a response to the research questions presented in chapter 1, and represent two sets of findings. The first set of findings concern the teacher comments, how they are presented to the students, what they focus on, and how the comments are framed. These findings are the response to the first research question that seeks to find out how teachers commented on student texts, and how their comments aligned with principles of formative assessment. The findings presented for this dataset show that while most of the teachers are employed at the same school, they have individual ways of providing feedback to student texts, though most focus on local text issues in the margin comments. The focus on local text issues supports the findings in Saliu-Abdulahi et.al (2017). The focus on form is more noticeable than other recent studies have suggested (Horverak, 2015, Bueie, 2016), and more in line with the pattern in many other L2 studies (Montgomery & Baker, 2007; Lee, 2008; Budimlic, 2012; Burner, 2015a). Those who commented at global levels mostly did so in the end comments, like in Hyland and Hyland (2001), and these comments varied greatly in length. Out of the 7 teachers in this study, 4 used the assessment criteria to explain how the student could improve their text. This supports the findings in Eriksen (2017), who found that roughly half of the teachers in his study provided feedback that is compliant with AfL principles. All teachers reported the process of assessing a decontextualized text to be a challenge, as in Solheim and Matre (2014).

The second set of data is a response to the second research question, which asks what types of feedback students prefer to receive on their texts and why certain types are preferred over others. This data set describes the focus group students' attitudes towards written feedback, both in general, and specifically based on the teacher comments. While some studies have found that students prefer comments directed at global levels (Bueie, 2016), and other studies have found that students prefer comments directed at local text levels (Lee, 2008; Saliu-Abdulahi, 2017), the students in this study did not show a clear preference for either when discussing comment types. In fact, students did not show as strong opinions towards any comment types as is reported in other studies (Straub, 2006; Bueie, 2016). Admittedly, the students preferred feedback that specified what was wrong, and how it could be fixed, though they felt it was unnecessary to receive all the answers, particularly regarding grammar and word choice issues. What was most important to students was that the

comments were specific, like in previous studies (Straub, 1997; Weaver, 2006; Burner, 2015a; Vågen, 2017), particularly when they provide advice and examples (Bueie, 2016). Furthermore, the students liked comments that were worded as questions rather than directives, because these make them think.

The reasoning for the focus group students' preferences may be found in their experiences with receiving feedback. The students assert that they seldom receive written feedback without grades, supporting findings in previous studies (Havnes et. al, 2012; Langseth, 2013; Nyvoll Bø, 2014; Saliu-Abdulahi, 2017). In fact, the students in this study felt that they received too little feedback, and that they wanted more feedback, confirming findings from yet another study (Lee, 2008). As a result of the feedback practices, the students in this study show a preference for feedback that is provided with grades. Moreover, they express that grades provide a sense of security, because they give an indication of how they are performing in a subject, like in Lipnevich and Smith (2009), though they do express that they do not always see the connection between the grades and the comments they receive. Still, the students state that when receiving satisfying grades, they do not necessarily read the feedback. The implications of the findings presented in this chapter will be discussed in the following chapter.

6. Discussion

The central purpose of this study was to explore the challenges of providing formative assessment in the intersection between assessment theory, teacher practice and student perceptions. Accordingly, it has been necessary to discuss and present what theory says about formative assessment, this was done in chapter 2. Furthermore, in chapter 3 an overview of the ‘state of affairs’ of how feedback is practiced and how students perceive these practices was provided. In chapter 5 the findings of this study were presented and compared to previous research. Extracted from the previous chapter, there are three main issues that merit discussion. These are that teachers tend to focus on local text levels, that teachers need context in assessment, and that the relationship between grades and feedback is problematic for students. These three issues point to more overarching concerns that will also be addressed, namely the relationship between formative and summative assessment, and assessment quality. In order to provide a lucid discussion of these issues, I will discuss the main issues in sections 6.1 through 6.3, before I discuss the overarching issues in sections 6.4 through 6.5.

6.1 The primacy of language

In this study, there was a strong tendency towards teachers commenting on language issues, such as grammar and vocabulary, illustrated in subsection 5.1.1. These findings are not surprising, considering that previous L2 research (see subsection 3.2.1) has indicated that language teachers tend to provide comments at local text levels. Other studies have suggested reasons ranging from external expectations, like school policy (Lee, 2008; Budimlic, 2012), to perceptions that students prefer this feedback (Budimlic, 2012), because these are easiest to correct (Saliu-Abdulahi, 2017). The latter argument is certainly in line with what Hoel (2000) says about unskilled writers, that they find feedback at local levels easiest to correct. While previous research has found various reasons for teachers’ focus on language in text assessment, I will discuss teachers’ emphasis on language drawing on theories of writing, views of learning and the English subject curriculum.

Writing is complex, it is a process that results in a product, and formative assessment must take both into account. In section 2.4, I presented two models of writing, ‘The Response Triangle’ and ‘The Wheel of Writing’, and explained how these models were useful for

discussing writing as a process and as a product. The strong emphasis on language in assessment indicates a view of writing as a product, because addressing language issues sends the signal that the text is complete, and all that is necessary is to correct language issues. Given the saliency of the strong language focus in this study, it is necessary to explore possible reasons for this.

Matre and Solheim (2014) suggest that teachers do not have sufficient knowledge about text levels, and that this results in a strong language focus in text assessment, because these are the least demanding to address. This could be the case, after all, as indicated in section 2.4, there is no single theory that can be used to explain what writing entails in all situations. The models presented in this thesis are two perspectives among many others. Knowing what writing entails depends on perspective. In other words, it is entirely possible that teachers know what it means to write, but that they emphasize the text as a finished product, and that their assessments reflect this. If revision is not presupposed, then there is no point in commenting on global text levels, both because it is more time-consuming, and because there is a real possibility that students ignore feedback when there is no time for revision, as indicated in section 5.3.

It should be noted though, that the teachers in this study were asked to provide feedback as if the text was a draft that would be revised (see appendix 2). Since the instructions implied that the teachers see the text as part of a process, one might have expected more comments aimed at developing the text, such as content and structure. It could be then, that the strong language focus indicates that teachers share an understanding of the primacy of linguistic correctness. The English subject curriculum emphasizes writing as a complex communicative skill that means adapting to the situation, and that writing is about acquiring knowledge and increasing understanding (see subsection 1.3.2). As previous research has indicated, students focus on improving what the teacher focuses on in the assessment. Consequently, the strong language focus is worrying, because it gives the impression that writing is about linguistic correctness, and nothing else. Since the instructions specifically asked that teachers take a process-oriented approach to the assessment, as formative feedback practices recommend (see sections 2.1 and 3.1), it is necessary to discuss how latent views of learning might have had an impact on the feedback focus in this study.

I have already pointed out how writing can be understood as both a process and a product. These two understandings of writing indicate different views of learning. Writing as a

process brings forth understandings of cognitive theories, discussed in subsections 2.2.2 and 2.4.1, and assessment should reflect this by stimulating thinking and development of higher order skills. Writing as a product can be seen in light of different views of learning. Consequently, assessment of writing as a product must be seen in light of different views of learning. Formative assessment is claimed to be situated within social theories of learning (see section 2.2). These theories emphasize that social interactions are crucial to the development of knowledge, and Vygotsky's 'zone of proximal development' is one example of this. The communicative function of writing, which is mentioned in the English subject curriculum (subsection 1.3.2), is an example of how social theories of learning are manifested in the curriculum. The fact that teachers are focusing their attention towards language issues and ignoring the more global text levels is interesting, considering that it is the global text levels that allow the students to demonstrate that they can adapt their writing to the situation and context, and that they can process information.

A strong focus on linguistic features that can be assessed as right or wrong can indicate a behaviorist perspective of text assessment. Language issues are easier to measure progress on, and while these are concrete, it is problematic to focus on these because they do not develop students' writing skills or thinking skills in the long run. Furthermore, this type of response may give the students the impression that it is the series of parts, the words, sentences and paragraphs, that are important, rather than what they are trying to convey, as Sommers worried in 1982 (see section 3.2). As this discussion has indicated, teachers ground formative assessment theoretically every time they teach and assess. Different views of learning are present in theories of writing, and in the curriculum. It is therefore important that teachers are aware of their own views of learning, and how these must be in alignment with the underlying theories in the curriculum, theories of writing, and in formative assessment. Basing assessment in one view of learning could ultimately result in ignoring important learning opportunities, and ignoring important skills that need to be learned.

Up to this point, much of the discussion has centered on the necessity of having knowledge of theoretical perspectives. It is therefore necessary to point out that teaching, and consequently assessing, means meeting demands of all sorts. Writing as a key competency is outlined in subsection 1.3.1, and is intended to be integrated into each subject, which means that assessment of writing has to take into consideration that each subject has its own understanding of knowledge, content, and how to demonstrate these. Assessing in any subject entails knowledge of how to encourage learning, motivation, progression towards

aims as well as meeting political demands. In other words, assessment competence is multi-faceted. Teachers need to have in-depth subject knowledge, so that they can gain a clear understanding of the aims that the students are working towards, as well as what progression towards these looks like. The Knowledge promotion reform has been criticized for the lack of a clear progression in competence aims. As long as there is a lack of a clear and unified view of how excellence in a subject area is defined, and what progression towards this looks like, there is bound to be implications for assessment practices. When a common understanding is lacking, teachers may seek out other ways to try to identify progression, such as including information about the student. This will be discussed in the following section.

6.2 Context as a factor in assessment

A few of the teachers in this study cited the lack of context as problematic for assessing the student text. Examples of some contextual factors they felt were lacking was information about the student's level, what content had been taught, and what the student's personal goals were. These findings correspond with findings in Solheim and Matre (2014, see subsection 3.2.2), though the researchers offer no explanation for this. In this section, I will discuss the perceived need for context from a theoretical writing perspective, and from a formative assessment perspective, which also means drawing on views of motivation and learning.

Solheim and Matre (2014) stated that some teachers found it difficult to assess the text as a text, and like some of the teachers in this study, found context to be necessary for assessment. The two models of writing discussed in section 2.4 have differing views of what a text is. 'The Response Triangle' focuses on the components of a text, whereas 'The Wheel of Writing' focuses on the acts of writing and the purposes these fulfil. In other words, both models of writing allow for reading a text without knowing about the student, if teachers are familiar with models of writing. It should therefore be possible, using the task wording and assessment criteria, to identify what the student needs to work more on to develop their writing.

The fact that teachers find context to be important is interesting from a formative assessment perspective. General principles of formative assessment do not cite the need for context. This is interesting on its own, because central to formative assessment is the idea of progress. It

therefore makes sense that teachers need context to provide assessment that promotes progress. Second, teachers' need for context indicates an understanding of formative assessment as part of a process. Ideally, formative assessment should be a loop, where a process leads to a product, that the teacher provides feedback to, and the student uses the feedback to improve. The previous section discussed why there was such a strong language focus. The lack of context could be one contributing factor. After all, if teachers lack contextual clues that provide them with an understanding of what the student needs to improve, it can be difficult to discern what is a simple mistake, and what is a recurring error.

The understanding that contextual factors are important is shared by students, who expect the teacher to tailor the feedback to their individual needs, as indicated in subsection 5.3.1. This shared understanding of the relevance of context could be based in a sociocultural understanding of formative assessment. In Vygotsky's 'zone of proximal development', the amount of support the teacher offers is in inverse proportion to the competence level of the student (subsection 2.2.3). This means that knowing what the student needs to work more on enables the teacher to tailor the assessment to the student's needs. Moreover, knowing what the student can master on their own is an advantage, because it saves time when providing feedback. Also, having contextual knowledge means knowing what the student finds motivating, which is an essential part of formative assessment. After all, if the student is not motivated to improve, no feedback will have this effect. In other words, knowing when to push the student towards improving and when to provide praise is easier if the teacher knows the student, and can make a difference in how the feedback is received. This is perhaps best illustrated by the fact that students stated that not all comments at local text levels need to be explicit. For the teacher to decide when the comment is explicit enough, context is helpful. As these two sections have tried to make clear, writing and formative assessment should ideally include elements of both product and process.

6.3 Assessment as reassurance

Unlike in previous studies (Straub, 1997; Bueie, 2016), the focus group student did not have clear preferences of feedback categories (see section 5.3). Admittedly, they preferred comments that asked questions, and that indicated where something was wrong, but not necessarily those that provided the correct answer. The students stated that they wanted comments that made them think. First of all, this is a positive finding in light of formative

assessment and the national curriculum, as both emphasize the importance of thinking about one's learning processes, a typical trait of cognitive theories. But more importantly, this indicates how crucial it is to not underestimate the student, and is a good example of teaching and assessing within the 'zone of proximal development'. The assessment should help the student close the gap between what they have managed and the criteria. The comments that make them think, whether they are questions or not, can in this way provide the scaffolding students need. In subsection 2.2.3, the 'zone of proximal development' was explained as providing the inverse amount of support to the students' competence level, and the students' preference of comments that make them think is an invitation to assess in this way. This reiterates why it is problematic that teachers tend to focus their attention on local text levels.

The more interesting findings from the focus group were those that explain why students lacked preferences of feedback types. First of all, students reported receiving little feedback, and were therefore very interested in receiving all types of feedback. Furthermore, students seldom received feedback without grades. The assessments that were presented without grades rarely included comments, but were often rubrics with highlighted sections or pass and fail. These assessment practices resulted in the students wanting grades, because they often felt unsure of what level their performances were at. The implication is that attempts made at assessing formatively without grades, were having the opposite effect than likely intended. Besides being unclear as to what the student had mastered, and what to work more on, this type of assessment did not help the student progress, because it did not provide any support in finding out *how* to improve, which is one of the principles of formative assessment. In other words, the students were not used to feedback practices that had the sole purpose of promoting learning. Students also stated that feedback was only provided post-product, resulting in a lack of revision work. The feedback practices resulted in something of a paradox that can be summed up as follows: the students wanted more feedback, but in combination with grades. Yet, if they received grades, they would not follow up on the feedback, particularly if the grade was satisfactory. However, they also claimed that if they only received feedback and no grades, they would use the feedback.

These findings show how formative assessment, motivation and grades are intertwined. In section 2.3, views of motivation are grounded in views of learning. Grades are typically extrinsic motivation, as described by Greeno et.al (1996), and they can have the effect of either increasing or decreasing motivation, as when the students in this study decide, based on their

grade, whether they do anything with the feedback or not. Like Huot and Perry (2009, see subsection 2.3.2) stated, grades seem to make the students work towards improving their grades rather than improving their writing skills. While it can be argued that these do not need to be separated, this depends on whether the feedback offers anything that can be used towards improving writing. It is doubtful that receiving feedback mainly on language issues develops writing skills.

In contrast, when students claim that they would use feedback if it was provided without grades, it could indicate that the students are intrinsically motivated to improve their work, and that grades have a negative impact on their motivation (Greeno et al, 1996, see section 2.3). If so, then this makes a strong argument for providing feedback without grades, as formative assessment practices should. Considering the claims made for the efficacy of formative assessment in chapter 2, one might wonder why teachers are still providing grades with feedback, particularly when also failing to provide students with time to make use of the feedback, since this makes the time spent on assessing wasted. At the very least, this points to the problematic relationship between formative and summative assessment, one of the overarching issues in this discussion.

6.4 A complicated relationship

The distinction between formative and summative assessment is complex and dependent on how one defines these. As discussed in subsection 1.4.2, it is indicated how the terms formative and summative tend to be used as denoting the timing of assessment. However, the distinction between formative and summative assessment is not that simple, and the same information can be used for summative or formative purposes. In other words, what the student and the teacher do with the assessment is what decides if the assessment is formative or not. In this study, the students' perceptions of the lack of correspondence between grades and feedback, outlined in subsection 5.3.2, is one finding that indicates that this relationship continues to be problematic. An example of this is that the feedback provided with grades contained too little information to see how the grades and feedback were expressions of the same result. For instance, one student had received one comment on word choice, and still did not receive the best grade.

Teachers are in a difficult position when they assess students' performances, they are accountable to school leaders, to students, and to political guidelines all at once. There is an

overall understanding that assessment should lead to learning, and also an understanding that this learning needs to be documented towards an end of year grade. This understanding is furthered by the fact that the Education Act (see subsection 1.3.3) was amended in 2015 to clarify the relationship between formative assessments during the year, and the end of term grade. The legislation specified that the competence documented throughout the year should count towards the end of year grade, effectively settling uncertainty as to how the formative assessments throughout the year fit into the bigger picture of summative accountability, but blurring the distinction between formative and summative assessments further. In other words, formative assessment is relevant for the final grade⁴, because it informs the teacher of the starting point, and what progress has been made. Consequently, quality formative assessment practices should make end of year grading more predictable for teachers and students.

The idea that progress during the year should count towards the end of year grade can be problematic. For one thing, the progress made during the year can be “wasted” in grading if the first assessments revealed a poor performance, as these can negatively affect the end of year performance. On the other hand, including the progress made could even out variable performances throughout the year. The second reason this might be a problem is if the formative assessments in the course of the year provide the students with too little information on how they are performing. It seems that this is the case for the students in this study, as they find the assessment practices unpredictable, which is why they prefer receiving grades. These provide a sense of “where” the student is performing according to the criteria. In other words, the findings indicate that the students feel the sense of how they are performing is lacking in the assessments they usually receive. As presented in subsection 5.3.2, the students in this study feel that the feedback practices vary at their school, and that they either receive feedback with grades, or rubrics with aim achievements that are highlighted, but without feedback.

From a student perspective this is troubling. If the teacher signals that this is a formative assessment situation, then the assessment the student receives should reflect this. For instance, when one teacher stated that he imagined the assessment situation as process-

⁴ Formative assessment is also relevant for the final exam, since the aim is improving the same skills throughout the year that the exam tests. Formative assessment is essentially working towards the same goal as the exam, but from a different direction.

oriented (section 5.2), which would fit well into a formative understanding of assessment, one might have expected that the assessment criteria would be referred to, though it was not. That the student is made aware of how they have performed against the criteria is crucial, according to generally accepted principles of formative assessment. Not providing this information is, essentially leaving the interpretation of the performance up to the student. Conversely, other teachers did refer to the assessment criteria to varying degrees, and signaled where the student could make improvements based on these, which seemingly is in line with formative assessment practices. However, the students report that they are not given the opportunity to work with the feedback, which begs the question of why they are shown possible areas of improvement. No matter how formative the intentions of an assessment are, they do not take on a formative function if students are not given the opportunity to use the feedback to improve their writing. In other words, if teachers do not provide students with the time to work on improvements after they have received feedback, then time spent providing feedback is time wasted for the teacher. It can be argued that the students can use the feedback to improve the next text, but this entails having students that are motivated to do so, and that are mature enough to see the need for it.

Assessing writing means that teachers need to have a good grasp of what writing is, and what it means to demonstrate writing competence in specific subjects. These notions of quality must be shared by other teachers if the assessment is to be considered fair and trustworthy, in order to limit the possibility of receiving feedback that is arbitrary. However, it is not necessarily easy to define a common understanding of what it means to demonstrate writing competence, illustrated by the fact that there are several models of writing, and that the curriculum leaves it to teachers to decide what progress and excellence look like. The result of a lack of a common understanding of what writing, progress and excellence mean result in assessment being based in individualized notions of writing competence. Consequently, the validity and reliability of the assessment is compromised, since it is based in individualized notions of quality.

6.5 Prerequisites for assessment quality

In section 3.1 the three conditions of formative feedback are outlined, that the students understand the goals they are aiming for, how they are performing compared to the goal, and what to do to reach the goal. This demands that the teacher has topic knowledge as well as

assessment competence. It has been indicated in subsection 2.2.5 how three principles are central to all forms of assessment, these are the principles that an assessment should be valid, reliable and transparent. An issue that needs to be discussed is how, and if, the principles that assessments should be valid, reliable, and transparent can apply to formative assessment, and what this means for the findings in section 5.2, that teachers found context to be necessary for providing feedback to students.

Hattie and Timperley (2007) state that formative assessment should answer three questions (section 3.1). The first of these is “where am I going” which indicates the convergence of assessment and transparency; that the student is aware of what is to be learned, and which criteria this will be judged by. Furthermore, it can be argued that before this question can be asked, another principle of assessment must first be addressed, such as validity. Assessment validity refers to the assessment form being appropriate to the assessment purpose, and the achievement of the assessment purpose (Stobart, 2012, see subsection 2.2.5), and is affected by the interpretations of curriculum that are made, that inform decisions which affect the assessment (Sandvik, et. al, 2012, see subsection 2.2.5).

Validity can be explained using the constructed assessment situation the teachers in this study faced. The purpose of the task is that the student demonstrates that they understand how English came to be a lingua franca, and how it can be a threat to other languages or cultures, this is clear in the task wording. The assessment form is that the student writes an essay, and the task clearly states what the student text is being measured by, and which aspects of writing and subject content these relate to. Since the criteria are clearly stated, the task can be said to have a high level of transparency, the student knows what they are aiming towards, and the teacher need only decide if the student text meets the criteria. In other words, before the assessment is designed, issues of validity are already in play. This is because validity depends on decisions of what the task is supposed to demonstrate, and how. This means that validity and transparency are important prerequisites for formative assessment as well as assessment in general. However, there is nothing in this understanding that opens up for the inclusion of context in the assessment situation, which begs the question of whether the call for contextual clues, as found in section 5.2, threatens the validity of the assessment.

Messick (1995) problematizes the common assumptions of validity, because validity, he claims, is more than the correlation between criteria and results, or the content and use of a

test (p. 742). He proposes that an understanding of validity should be expanded to include understandings of the social consequences of how the assessment is used, as well as the values that underlie the meanings of assessment (p. 741). This understanding allows for including context as a factor in assessing, and emphasizes how the actions taken on the basis of the inferences, which are made as a result of assessing, has an impact on learning and motivation. In other words, the interpretations that are made on the basis of an assessment can be strengthened, or challenged, by knowing the student, and in this way increase the fairness of the assessment. The fairness of the assessment is particularly important, considering that progress can count towards end of year grades.

It can be argued that context is irrelevant for providing the student with a fair assessment. After all, as long as one has criteria, and an understanding of what constitutes achievement of these criteria, then, according to general principles of feedback, all the tools are in place to assess the students. Furthermore, there should be enough information to inform the students how to proceed, based on their performance. However, since part of the validity of assessment rests on the inferences made based on the results (see subsection 2.2.5), the inclusion of context can strengthen the validity of these. The knowledge we have of what the student knows before the assessment can help to reinforce our perceptions of what the student is capable of, what they need to work more on, and how they can do towards improving. If the perceived importance of context is a finding that is supported in later studies, then this issue questions the commonly accepted notion that general principles of feedback can be used to provide feedback regardless of context.

Reliability in formative assessment is also a problematic issue, as reliability often carries notions of objectivity (see subsection 2.2.5). The problem of reliability in formative assessment is evident in the findings presented in section 5.2, that teachers found the task of assessing the student text to be difficult, citing a lack of context as the reason. The notions of objectivity are particularly ill-suited for the type of assessment situation the teachers in this study faced, since this entails making qualitative judgements. In the example of the constructed situation, the assessment must measure whether the student has shown that they understand, which means analyzing whether the student has demonstrated the skills and knowledge in the task. Judgements of whether this has been achieved depends on to which degree the teacher finds the text convincing, rendering the assessment in part a matter of individual judgement. As this demonstrates, reliability in formative assessment may be difficult to achieve, as it depends on eliminating disturbances that may affect the assessment,

such as who is assessing the text, as explained by Black and Wiliam, 2012 (see subsection 2.2.5). This is because there is bound to be variations due to individual opinions of what a good text contains, which is supported by the finding in subsections 5.1.1 and 5.1.2, that nearly all teachers in this study had individual ways of commenting on the text. While this can be partly alleviated through a common understanding of constitutes quality in writing, formative assessment must take progress as well as the finished product into account. Progress cannot easily be quantified, and reliability should ensure consistency of results, quite the opposite of progress.

What this discussion means is that reliability, as it is commonly understood, may not be an appropriate factor to take into consideration in the type of task the teachers in this study were asked to assess. It might be necessary, then, for an expanded understanding of what reliability means in formative assessment. In subsection 2.2.5, a reconceptualization of reliability is offered. Smith (2003) terms this ‘sufficiency of information’, and describes this as the teacher having enough information to decide how the student performance compares to the aim. This entails asking questions of validity, such as if the content being assessed is representative of the subject, and if the task will provide sufficient information about the students’ learning. In the case of the latter, context can be quite useful, as it can expand on the information the teacher needs.

As this discussion has indicated, formative assessment is complex. It entails being aware of one’s own views of learning, how formative assessment is situated within more than one view of learning, and how these must both be in alignment with the curriculum. Formative assessment of writing is no less complex, as it demands knowledge of what writing is, both the finished product, and the process of getting there. Teachers must also know what it means to be motivated, and how to stimulate this in students. In addition, teachers must have a solid foundation built on subject knowledge. The sum of these contribute to a teacher’s assessment competence. This assessment competence is instrumental in securing the quality of assessment, demonstrated through transparency, validity and sufficiency of information, and ultimately in securing the quality of students’ education and preparation towards exams. The sum of this indicates that assessment is complex, and as such, demands a great deal of teachers.

7. Concluding remarks

This thesis has used written teacher feedback, and how it is received by students, to discuss possible challenges of formative assessment. To this end, three research questions have served to limit the scope of this thesis, and have guided this study. The first research question asked what formative assessment theory claims best assessment practice is. To approach this question, I presented general principles of formative assessment, and discussed how views of learning could serve as theoretical bases. The answer is a new question, and that is if general principles can apply to all subjects and subject areas. In light of the findings here and in previous research there is reason to believe that formative assessment practices could be improved by first defining what constitutes important knowledge in the English subject, as this will help create a common understanding of what to focus on in assessment.

The second research question asked what types of written feedback teachers give students on texts, and what they say about this. The answer is that they mainly give feedback at the local text levels, and that they feel context is necessary when providing feedback. This raises the question of the implications this has for the fairness of assessment, and whether contextual factors improve, or detract from the quality of the assessment. The finding that context is considered to be important for formative assessment is something that needs to be further investigated, to find out why teachers need context, and what this says about their understandings of formative assessment.

The third research question asked what students' perceptions of feedback practices are. In short, the students are motivated by feedback that allows them to think on their own. Furthermore, the quality, amount, and timing of feedback have an impact on if they use it for improving their text. Interestingly, summative assessment purposes infringe on students' willingness to accept purely formative feedback. Yet, they want more formative feedback, but in conjunction with grades. Herein lies a paradox, because while formative assessment must have a goal, this goal seems to have become more important than how to get there. In light of this, there is a clear necessity of discussing feedback practices, the relationship between formative and summative assessment, and how this relationship affects what is intended to be purely formative feedback.

All of these questions and discussions have contributed to a wider discussion, where the challenges of providing formative feedback has been the primary issue. This discussion,

while not providing answers, has problematized the general principles of formative assessment. There are multiple problems, the first being the notion that general principles can apply to all subjects and skills. This means that there is an understanding that the same principles of assessing can apply to math and language, and is surprising, as one would never consider the same teaching practice to apply for both. Secondly, the intentions of formative assessment in a summative climate seem to impact assessment practices negatively, in that teachers seldom provide purely formative assessment, and students therefore prefer grades with their feedback. Finally, the preconditions of validity and reliability in assessment are problematized, particularly since there does not seem to be an understanding of reliability that can be used in conjunction with how formative assessment is currently defined. All of these discussions point to a need for defining what formative assessment means in the English subject, and what it means for writing within the English subject. This is a necessary precondition for improving assessment practices.

Reference list

Assessment Reform Group (2002) *Assessment for Learning: 10 principles research-based principles to guide classroom practice*. London: Assessment Reform Group.

Bachman, L. F. (2004). *Statistical analyses for language assessment*. Cambridge: Cambridge University Press.

Bachman, L. & Palmer, A. (2010). *Language assessment in practice*. Oxford: Oxford University Press.

Bennet, R. E. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5-25.

Berge, K. L., & Skar, G. (2015). *Ble elevene bedre skrivere? Intervensjonseffekter på elevers skriveferdigheter og skriveutvikling*. (NORM Report 2, 2015). Høgskolen i Sør-Øst Trøndelag. Retrieved from http://norm.skripesenteret.no/wp-content/uploads/2015/08/Berge_og_Skar-2015-NORM-rapport2.pdf

Berge, K. L., Evensen, L. S., & Thygesen, R. (2016). The Wheel of Writing: a model of the writing domain for the teaching and assessing of writing as a key competency. *The Curriculum Journal*, 27(2), 172-189. <http://dx.doi.org/10.1080/09585176.2015.1129980>

Black, P., & Wiliam, D. (1998a). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy and Practice*, 5(1), 7-74. <http://dx.doi.org/10.1080/0969595980050102>

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Maidenhead: Open University Press.

Black, P., & Wiliam, D. (2012). The Reliability of Assessments. In J. Gardner (Ed.) *Assessment and Learning* (2nd ed.), (pp. 243-263). London: SAGE.

Brannon, L., & Knoblauch, C. H. (1982). On Students' Rights to Their Own Texts: A Model of Teacher Response. *College Composition and Communication*, 33(2), 157-166. <http://dx.doi.10.2307/357623>

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <http://dx.doi.org/10.1191/1478088706qp063oa>

Brookhart, S. M. (2003). Developing Measurement Theory for Classroom Assessment Purposes and Uses. *Educational Measurement: Issues and Practice*, 22(4), 5-12. <http://dx.doi.org/10.1111/j.1745-3992.2003.tb00139.x>

Brorsson, B. N. (2007). Man liksom bara skriver. Skrivande och skrivkontexter i grundskolans år 7 och 8. (Doctoral dissertation). Retrieved from <http://www.diva-portal.org/smash/get/diva2:137474/FULLTEXT01.pdf>

Budimlic, D. (2012). *Written feedback in English. Teacher's Practices and Cognition*. (Master's thesis). Norwegian University of Science and Technology, Trondheim. Retrieved from <https://brage.bibsys.no/xmlui/handle/11250/2460792>

Bueie, A. A. (2016). Nyttige og mindre nyttige lærerkommentarer – slik elevene ser det. *Nordic Journal of Literacy Research*, 2, 1-28. <http://dx.doi.org/10.17585/njlr.v2.188>

Burner, T. (2015a). Formative assessment of writing in English as a foreign language. *Scandinavian Journal of Educational Research*, 60(6), 626-648. <http://dx.doi.org/10.1080/00313831.2015.1066430>

Burner, T. (2015b). Processes of change when using portfolios to enhance formative assessment of writing. *Assessment Matters*, 9(15), 53-79. <http://dx.doi.org/10.18296/am.0011>

Bø, E. N. (2014). *A case study of feedback to written English in a Norwegian upper secondary school*. (Master's thesis). University of Stavanger, Stavanger. Retrieved from <http://brage.bibsys.no/xmlui/handle/11250/198753>

Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education*. (4th ed.) London: Routledge.

Creswell, J. W. (2007). *Qualitative inquiry & research design: choosing among five approaches*. (2nd ed.) Thousand Oaks, Calif: SAGE.

Dale, E. L., Engelsen, B. U., & Karseth, B. (2011). *Kunnskapsløftets intensjoner, forutsetninger og operasjonaliseringer: En analyse av en læreplanreform*. Oslo: University

of Oslo. Retrieved from
http://www.udir.no/Upload/Rapporter/2011/5/PFI_sluttrapport_2011.pdf?epslanguage=no

Denscombe, M. (2014). *The good research guide. For small-scale social research projects.* (5th ed.). Maidenhead: Open University Press.

Dysthe, O. (2008). Klasseromsvurdering og læring. *Bedre Skole*, (4), 16-23. Retrieved from:
<https://www.udir.no/globalassets/filer/vurdering/vfl/andre-dokumenter/felles/olga-dyste-bedre-skole-08.pdf>

Dysthe, O. (2009). Læringssyn og vurderingspraksis. In J. Frost (Ed.) *Evaluering- i et dialogisk perspektiv.* (pp.33-52). Oslo: Cappelen Damm.

Dysthe, O., & Hertzberg, F. (2009). Den nyttige tekstresponsen–hva sier nyere forskning? In O.K. Haugaløkken, L.S. Evensen, F. Hertzberg & H. Otnes (Eds.). *Tekstvurdering som didaktisk utfordring* (pp. 35-43). Oslo: Universitetsforlaget.

Ellis, R. (2009). Corrective Feedback and Teacher Development. *L2 Journal*, 1(1), 3-18.

Engh, K. R., Dobson, S. & Høihilder, E. K. (2007). *Vurdering for læring.* Kristiansand: Høyskoleforlaget.

Eriksen, H. (2017). Vurdering for læring i norskfaget: Hva kjennetegner norsklæreres skriftlige tilbakemeldinger til egne elever? *Acta Didactica Norge*, 11(1), 1-26. <http://dx.doi.org/10.5617/adno.4078>

Evensen, L. S. (2009). Vurdering av skrivekompetanse: En kompleks utfordring. In O. K. Haugaløkken, L. S. Evensen, F. Hertzberg & H. Otnes (Eds.). *Tekstvurdering som didaktisk utfordring* (pp.15-23). Oslo: Universitetsforlaget.

Ferris, D. R. (2014). Responding to student writing: Teachers' philosophies and practices. *Assessing Writing*, 19, 6-23. <https://dx.doi.org/10.1016/j.asw.2013.09.004>

Fjørtoft, H. (2013). Vurderingspraksiser i norskfaget. In Sandvik, L. V., Buland, T. (Eds.), *Perspektiver på individuell vurdering i skolen.* Retrieved from:
<http://www.udir.no/Upload/Rapporter/2013/FIVIS2.pdf?epslanguage=n>

Forskrift til opplæringslova. (2009). Retrieved from https://lovdata.no/dokument/SF/forskrift/2006-06-23-724/KAPITTEL_4#KAPITTEL_4

Gamlem, S. M. & Smith, K. (2013). Student perceptions of classroom feedback. *Assessment in Education: Principles, Policy & Practice*, 20(2), 150-169. <http://dx.doi.org/10.1080/0969594X.2012.749212>

Gardner, J. (2012). Assessment and Learning: Introduction. In J. Gardner (Ed.), *Assessment and Learning* (2nd ed.), (pp.1-8). London: SAGE.

Grabe, W. & Kaplan, R. B. (1996). *Theory & Practice of Writing*. London: Longman.

Graham, S. & Weiner, B. (1996). Theories and Principles of Motivation. In D. C. Berliner & R. C. Calfee. (Eds.), *Handbook of Educational Psychology* (pp. 63-84). NY: Simon & Schuster Macmillan.

Greeno, J. G., Collins, A. M. & Resnick, L. (1996). Cognition and Learning. In D.C. Berliner & R.C. Calfee. (Eds.), *Handbook of Educational Psychology* (pp. 15-46). NY: Simon & Schuster Macmillan.

Gröndahl, K. P. (2015). *EFL Pupils' Understanding and Use of Teacher Written Feedback. A case study in two Swedish lower secondary classrooms*. (Licentiate Thesis). Stockholm University.

Halkier, B. (2010). *Fokusgrupper*. Oslo: Gyldendal Akademisk.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <http://dx.doi.org/10.3102%2F003465430298487>

Harlen, W. (2012). On the Relationship Between Assessment for Formative and Summative Purposes. In J. Gardner (Ed). *Assessment and Learning* (2nd ed.), (pp. 87-102). London: SAGE.

Havnes, A., Smith, K., Dysthe, O., Ludvigsen, K. (2012). Formative assessment and feedback: Making learning visible. *Studies in Educational Evaluation*, 38(1), 21-27. <https://dx.doi.org/10.1016/j.stueduc.2012.04.001>

Hoel, T. L. (2000). *Skrive og samtale: Responsgrupper som læringsfellesskap*. Oslo: Gyldendal Akademisk.

Horverak, M. O. (2015). Feedback practices in English in Norwegian upper secondary schools. *Norwegian Journal of Modern Language Methodology*, 3(2), 74-9. Retrieved from <http://journal.uia.no/index.php/NJMLM/article/view/140/224#.WnRUkKjiY2w>

Horverak, M. O. (2016). A survey of students' perceptions of how English writing instruction is carried out in Norwegian upper secondary schools. *Journal of Second Language Teaching and Research*, 5(1), 120-144. Retrieved from <http://pops.uclan.ac.uk/index.php/jsltr/index>

Huot, B. & Perry, J. (2009). Toward a New Understanding for Classroom Writing Assessment. In R. Beard, D. Myhill, J. Riley and M. Nystrand (Eds.), *The SAGE Handbook of Writing Development* (pp. 423- 435). London: SAGE.

Hyland, F. & Hyland, K. (2001). Sugaring the pill: Praise and criticism in written feedback. *Journal of second language writing*, 10(3), s. 185–212.

Hyland, K. (2003). *Second Language Writing*. Cambridge: Cambridge University Press.

Igland, M.-A. (2008). *Mens teksten blir til. Ein kasusstudie av lærarkommentarar til utkast*. (Doctoral dissertation). University of Oslo.

Kronholm-Cederberg, A. (2009). *Skolans responskultur som skriftpraktik. Gymnasisters berättelser om lärarens skriftliga respons på uppsatsen*. (Doctoral dissertation). Åbo Academy University Press.

Kvithyld, T. & Aasen, A. J. (2011). Fem teser om funksjonell respons på elevtekster. *Tidsskriftet Viden om Læsning*, (9), 10-16. Retrieved from http://www.skrivesenteret.no/uploads/files/Trygve_Kvithyld_Arne_Johannes_Aasen1.pdf

Kvale, S., Brinkman, S. (2009). *Det kvalitative forskningsintervju* (2.utg). Oslo: Gyldendal Akademisk.

Langseth, I. (2013). Vurderingspraksiser i engelskfaget. In Sandvik, L. V., Buland, T. (Eds.), *Perspektiver på individuell vurdering i skolen*. Retrieved from <http://www.udir.no/Upload/Rapporter/2013/FIVIS2.pdf?epslanguage=n>

Lee, I. (2008). Student reactions to teacher feedback in two Hong Kong secondary classrooms. *Journal of Second Language Writing*, 17(3), 144-164. <https://dx.doi.org/10.1016/j.jslw.2007.12.001>

Lincoln, Y. S. & Guba, E. G. (1985). *Naturalistic Inquiry*. Newbury Park: SAGE.

Lipnevitch, A. A. & Smith, J. K. (2009). "I really need feedback to learn:" students' perspectives on the effectiveness of the differential feedback messages. *Educational Assessment, Evaluation and Accountability*, 21(4), 347-367. <http://dx.doi:10.1007/s11092-009-9082-2>

Matre, S. & Solheim, R. (2014). Lærersamtalar om elevtekstar – mot eit felles fagspråk om skriving og vurdering. In *Alle tiders norskdidaktiker. Festskrift til Frøydis Hertzberg på 70-årsdagen*. (pp. 219-243). <http://dx.doi.org/10.1037/0003-066X.50.9.741>

Messick, S. (1995). Validity of Psychological Assessment. *American Psychologist*, 50(9), 741-774. <http://dx.doi.org/10.1037/0003-066X.50.9.741>

Moeed, A. (2015). Theorizing Formative Assessment: Time for a Change in Thinking. *The Educational Forum*, 79(2), 180-189. <http://dx.doi.org/10.1080/00131725.2014.1002593>

Montgomery, J. L., & Baker, W. (2007). Teacher-written feedback: Student perceptions, teacher self-assessment, and actual teacher performance. *Journal of Second Language Writing*, 16(2), 82-99. <http://dx.doi.org/10.1016/j.jslw.2007.04.002>

Norwegian Directorate for Education and Training. (2012). *Framework for basic skills*. Retrieved from https://www.udir.no/contentassets/fd2d6bfbf2364e1c98b73e030119bd38/framework_for_basic_skills.pdf

Norwegian Directorate for Education and Training. (2013). *English Subject Curriculum*. Retrieved from <https://www.udir.no/kl06/ENG1-03>

Norwegian Directorate for Education and Training. (2015a). *Vurdering for læring – om satstingen*. Retrieved from <https://www.udir.no/laring-og-trivsel/vurdering/nasjonalsatsing/om-satsingen/>

Norwegian Directorate for Education and Training. (2015b). *Fire prinsipper for god undervisvurdering*. Retrieved from <https://www.udir.no/laring-og-trivsel/vurdering/om-vurdering/undervisvurdering/>

Norwegian Ministry of Education and Research. (2007). *Early Intervention for Lifelong Learning 2006-2007*. Retrieved from <https://www.regjeringen.no/en/dokumenter/report-no.-16-to-the-storting-2006-2007/id441395/>

Norwegian Ministry of Education and Research. (2016). *Fag—Fordypning—Forståelse: En fornyelse av Kunnskapsløftet 2015-2016*. Retrieved from <https://www.regjeringen.no/no/dokumenter/meld.-st.-28-20152016/id2483955/?q=fag%20fordypning%20forst%C3%A5else>

Nusche, D., Earl, L., Maxwell, W., & Shewbridge, C. (2011). *OECD Reviews of Evaluation and Assessment in Education: Norway*. Paris: OECD publishing.

Patton, M. Q. (1990). *Qualitative evaluation and research methods* (2nd ed.). Newbury park: SAGE.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral Science*, 28(1), 4-13. <http://dx.doi.org/10.1002/bs.3830280103>

Sadler, R. D. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119-144. <http://dx.doi.org/10.1007/BF00117714>

Saliu-Abdulah, D., Hellekjær, G. O., Hertzberg, F. (2017). Teachers' (formative) feedback practices in EFL writing classes in Norway. *Journal of Response to Writing*, 3(1), 31-55. Retrieved from <http://www.journalrw.org/index.php/jrw/article/view/69>

Saliu-Abdulah, D. (2017). Scaffolding writing development: How formative is the feedback? *Moderna Språk*, 111(1), 127-155. Retrieved from <http://ojs.ub.gu.se/ojs/index.php/modernasprak/article/view/3541/3174>

Sandvik, L.V. (2011). *Via mål til mening: en studie av skriving og vurderingskultur i grunnskolenes tyskundervisning*. (Doctoral thesis). Norwegian university of Science and Technology: Trondheim. Retrieved from <http://hdl.handle.net/11250/270227>

Sandvik, L.V., Engvik, G., Fjørtoft, H., Langseth, I. D., Aaslid, B.E., Mordal, S & Buland, T. (2012). *Vurdering i skolen. Intensjoner og forståelser*. Retrieved from <http://www.udir.no/Upload/Rapporter/2012/fivis.pdf?epslanguage=no>

Sfard, A. (1998). On Two Metaphors for Learning and the Dangers of Choosing Just One. *Educational Researcher*, 27(2), 4-13. <http://dx.doi.org/10.2307/1176193>

Shepard, L. A. (2000). The Role of Assessment in a Learning Culture. *Educational Researcher*, 29(7), 4-14. <http://dx.doi.org/10.3102/0013189X029007004>

Smith, J. K. (2003). Reconsidering Reliability in Classroom Assessment and Grading. *Educational Measurement: Issues and Practice*, 22(4) 26-33. <http://dx.doi.org/10.1111/j.1745-3992.2003.tb00141.x>

Solheim, R., & Matre, S. (2014). Forventninger om skrivekompetanse. Perspektiver på skriving, skriveopplæring og vurdering i «Normprosjektet». *Tidsskriftet Viden om Læsning* (15), 76-89. Retrieved from http://norm.skrivesenteret.no/wp-content/uploads/2014/03/videnom_15_8.pdf

Sommers, N. (1982). Responding to Student Writing. *College Composition and Communication*, (33)2, 148-156. Retrieved from <http://www.jstor.org/stable/357622>

Stobart, G. (2008). *Testing Times: The uses and abuses of assessment*. New York/London: Routledge.

Stobart, G. (2012). Validity in Formative Assessment. In J. Gardner (Ed.). *Assessment and Learning* (2nd ed.), (pp. 223-242). London: SAGE.

Straub, R. (1997). Students' reactions to teacher comments: An exploratory study. *Research in the Teaching of English*, p. 91–119. Retrieved from http://www.jstor.org.ezproxy.hihm.no/stable/40171265?seq=1#page_scan_tab_contents

Taras, M. (2010). Assessment for learning: assessing the theory and evidence. *Procedia Social and Behavioral Sciences*, 2(2), 3015-3022. <http://dx.doi.org/10.1016/j.sbspro.2010.03.457>

Thronsen, I., Hopfenbeck, T.N., Lie, S., Dale, E.L. (2009). *Bedre vurdering for læring: Rapport fra "Evaluering av modeller for kjennetegn på måloppnåelse i fag"*. Oslo:

University of Oslo. Retrieved from https://www.udir.no/globalassets/upload/forskning/5/bedre_vurderingspraksis_ils_rapport.pdf

Thygesen, R., Berge, K. L., Evensen, L. S., Fasting, R. B. (2007). Sluttrapport: *Nasjonale prøver i skrivning som grunnleggende ferdighet*. Nasjonalt senter for leseopplæring og leseforskning: Universitetet i Stavanger.

Torrance, H. (1993). Formative assessment: Some theoretical problems and empirical questions. *Cambridge Journal of Education*, 23(3), 333-343.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46(2), 327-369.

Vik, L. B. (2013). *The assessment of English in two Norwegian upper secondary schools: A comparative study*. (Master's thesis). University of Stavanger. Retrieved from <http://brage.bibsys.no/xmlui/handle/11250/198753>

Vågen, M. T. A (2017). *Formative assessment in EFL writing: A case study of pupils' perceptions of their feedback practice and attitudes to receiving and using feedback*. (Master's thesis). University of Bergen. Retrieved from <http://bora.uib.no/bitstream/handle/1956/15942/MASTERny.pdf?sequence=1>

Weaver, M. R. (2006). Do Students Value Feedback? Student Perceptions of Tutors' Written Responses. *Assessment & Evaluation in Higher Education* 31(3), 379-394. <http://dx.doi.org/10.1080/02602930500353061>

Woolfolk, A. E. (2001). *Educational Psychology* (8th ed.) Boston: Allyn & Bacon.

Appendices

Appendix 1- Approval from NSD



Christina Sandhaug

2418 ELVERUM

Vår dato: 23.11.2017

Vår ref: 57236 / 3 / HUT

Deres dato:

Deres ref:

Vurdering fra NSD Personvernombudet for forskning § 31

Personvernombudet for forskning viser til meldeskjema mottatt 16.11.2017 for prosjektet:

57236	<i>Student attitudes toward the usefulness of written feedback – a study of high school students in one county in Norway</i>
Behandlingsansvarlig	<i>Høgskolen i Innlandet, ved institusjonens øverste leder</i>
Daglig ansvarlig	<i>Christina Sandhaug</i>
Student	<i>Catharine Meissner</i>

Vurdering

Etter gjennomgang av opplysningene i meldeskjemaet og øvrig dokumentasjon finner vi at prosjektet er meldepliktig og at personopplysningene som blir samlet inn i dette prosjektet er regulert av personopplysningsloven § 31. På den neste siden er vår vurdering av prosjektopplegget slik det er meldt til oss. Du kan nå gå i gang med å behandle personopplysninger.

Vilkår for vår anbefaling

Vår anbefaling forutsetter at du gjennomfører prosjektet i tråd med:

- opplysningene gitt i meldeskjemaet og øvrig dokumentasjon
- vår prosjektvurdering, se side 2
- eventuell korrespondanse med oss

Vi forutsetter at du ikke innhenter sensitive personopplysninger.

Meld fra hvis du gjør vesentlige endringer i prosjektet

Dersom prosjektet endrer seg, kan det være nødvendig å sende inn endringsmelding. På våre nettsider finner du svar på hvilke [endringer](#) du må melde, samt endringskjema.

Opplysninger om prosjektet blir lagt ut på våre nettsider og i Meldingsarkivet

Vi har lagt ut opplysninger om prosjektet på nettsidene våre. Alle våre institusjoner har også tilgang til egne prosjekter i [Meldingsarkivet](#).

Dokumentet er elektronisk produsert og godkjent ved NSDs rutiner for elektronisk godkjenning.

Appendix 2 Instructions for teachers

Survey of student's preferences of written teacher feedback on written texts in English

Thank you for your willingness to participate in my master's project. The aim of this project is to find out what types of written feedback teachers give, and what types students prefer. Your comments and feedback on this text will be made fully anonymous.

Instructions:

The text included in this document is written by a student in VG1 in the general studies program as part of a course in process writing. This text is the first draft, and the student is expecting feedback in order to revise their text. With this in mind, give feedback as you normally would. **You do not need to give the text a grade.** The task description and assessment criteria are included below, and were made available to the student before they began writing.

Topic: English as a global language

Task:

Write an essay in which you address the following questions:

- How did English become the most powerful language in the world?
- In what ways is English the Lingua Franca of the world?
- Could English be a threat to other languages or cultures?

Assessment criteria	Vurderingskriterier
Content: <ul style="list-style-type: none"> - You answer the questions raised in the task - You explain issues related to the topic and terms - You use examples, discuss from different viewpoints, point to causes and effects, and give grounds for your opinions - You use sources independently and critically, and cite these 	Innhold: <ul style="list-style-type: none"> - Du svarer på det oppgaven ber om - Du forklarer begreper og problemstillinger relatert til temaet - Du bruker eksempler, vurderer for og imot, peker på årsaker og konsekvenser, argumenterer for egne meninger - Du viser selvstendighet og evnen til å være kritisk i omgang med kilder, samt at du refererer til kildene du bruker
Structure: <ul style="list-style-type: none"> - Your text has a suitable title - Your text has an introduction that presents the topic - The main body of your text addresses the topic in a logical fashion - Your text has a conclusion that sums up the main points from your text 	Struktur: <ul style="list-style-type: none"> - Teksten har en egnet overskrift - Teksten har en innledning som presenterer emnet - Teksten har en hoveddel der du fordypet deg i temaet i en logisk rekkefølge - Teksten har en avslutning med oppsummering/konklusjon
Language: <ul style="list-style-type: none"> - You use varied vocabulary - You use formal language and express your thoughts in a precise manner - You use varied sentence structures (meaning that you don't start all sentences the same), and have correct grammar, spelling and punctuation - Your English is idiomatic (means that your English is fluent) 	Språk: <ul style="list-style-type: none"> - Du bruker et variert ordforråd - Du bruker et saklig og presist språk - Du varierer setningsoppbygningen (betyr at du ikke begynner alle setninger likt), og mestrer rettskriving og tegnsetting, samt grammatikk - Du skriver idiomatisk engelsk (betyr at du skriver engelsk med flyt)

Appendix 3 sample of student text with inserted comments

How did English become a world language?

What does "the empire on which the sun never sets" means? This phrase was used to describe the British Empire in the end of 19th centuries. The empire was so big that covered 35.500.000 square kilometers. Around 1/4 of total world population lived in the British empire, and used English as an official language to communicate. However, the Empire doesn't exist anymore. But language and culture the British empire had [left] can still be seen in every [land] where the empire had ruled.

In the end of 17th centuries the British navy successfully defeated his enemies, and became a strong power over the sea. The empire [grew] fast because of that, and became to the largest colony empire in the history. Colonization helped the spread of English to the world. The [industry] revolution gave the British empire technology, and goods from colonies became important for production in the Great Britain. [Handle] between colonies and Great Britain became more than before. [Many] British people moved from the Great Britain to colonies to discover a new life. Those immigrants brought the English culture and language from the Great Britain to colonies, and spread it all around the world.

English is important to make a conversation with people on internet, or people who cannot speak your language. It has become a "Lingua Franca" of the world. [There is] many reasons for why English is a Lingua Franca. The first reason is a lot of people have English as their mother tongue, and English is also the biggest second language in the world. [Kids getting English lessons on the school even they is] not living in an English-speaking country. [The other reason is that English is very useful in our daily life. Social media is getting stronger. More and more people begin to use programs like Facebook, Twitter, snapchat and Instagram. Distance is not a problem to communicate. People can communicate even they is thousands of kilometers away from each other. English is a useful language when people wants a conversation with another [guy] who have different [language on internet]. For example, in handle between countries or chat with a foreigner.

- Catharine Vero... C1 Your title addresses one side
- Catharine Vero... C2 Interesting intro. It makes me
- Catharine Vero... C3 Can you somehow bring up?
- Catharine Vero... C4 behind
- Catharine Vero... C5 A better word would be
- Catharine Vero... C6 This is a good sentence to
- Catharine Vero... C7 Take a second look at the
- Catharine Vero... C8 Which enemies?
- Catharine Vero... C9 to grow – grew – has grown?
- Catharine Vero... C10 Industrial
- Catharine Vero... C11 Handle means to take care
- Catharine Vero... C12 This paragraph nicely sums
- Catharine Vero... C13 Switch these paragraphs,
- Catharine Vero... C14 Could you use a linking
- Catharine Vero... C15 There is one reason, but
- Catharine Vero... C16 at
- Catharine Vero... C17 Rephrase this sentence
- Catharine Vero... C18 though
- Catharine Vero... C19 I am' you are /he she it is'
- Catharine Vero... C20 You present both the
- Catharine Vero... C21 person
- Catharine Vero... C22 Try to cover other uses of'

Appendix 4 Information for students and parents

Forespørsel om å delta i masterprosjekt- Informasjon og samtykkeskjema

Prosjektansvarlige:

Christina Sandhaug

Catharine Meissner

I forbindelse med min masterutdanning ved Høgskolen i Innlandet gjennomfører jeg et prosjekt det jeg skal undersøke hvordan elever forstår lærerens kommentarer på tekster elevene skriver i engelskfaget. Elever skriver en del tekster i engelskfaget, og engelsklærere bruker mye tid på å gi tilbakemelding på slike tekster. Vi har noe kunnskap om tilbakemeldingenes effekt, men vi vet lite om hvordan elevene forstår og hvordan de klarer å nyttiggjøre seg tilbakemeldingene de får. Hensikten med mitt prosjekt er å undersøke nettopp dette. Målet med prosjektet er å bidra til økt forståelse for hvordan elever forstår og nyttiggjør seg lærerens tilbakemeldinger.

Forespørselen om deltakelse i dette prosjektet går ut til elever i en VG1 klasser ved en videregående skole i Hedmark. Deltakelse i prosjektet innebærer at elevene vil delta i en fokusgruppe på 4 elever der de skal diskutere og vurdere nyttigheten av lærerkommentarer. Deltakere blir valgt ut i samråd med trinnets engelsklærere. Intervjuene vil bli tatt opp med lydopptaker. I fokusgruppen blir elevene bedt om å sette ord på hvordan de forstår lærerens kommentarer og hvilke tanker elevene gjør seg om disse kommentarene.

Det er helt frivillig å delta i prosjektet, og deltakerne kan på hvilket som helst tidspunkt trekke seg fra prosjektet uten å måtte begrunne dette nærmere. I så tilfelle blir datamaterialet slettet. Opplysningene som samles inn, vil bli behandlet konfidensielt. Resultatene av studien vil bli publisert slik at den enkelte ikke kan gjenkjennes. Mastergradsarbeidet forventes å være avsluttet våren 2017. Etter at prosjektet er avsluttet, vil opplysningene bli anonymisert.

Dersom du ønsker å delta i undersøkelsen, er det fint om du signerer den vedlagte samtykkeerklæringen og returnerer den til engelsklærer så snart som mulig. Både elev og foresatte må undertegne samtykkeerklæringen.

Har du spørsmål i forbindelse med denne henvendelsen, kan du gjerne ta kontakt med meg på e-post: catmeis@gmail.com eller telefonnummer 45867103.

Med vennlig hilsen

Catharine Meissner

Samtykkeerklæring:

Jeg har mottatt informasjon om prosjektet, og er villig til å delta i studien.

Elevens signatur:

Jeg har mottatt informasjon om prosjektet, og samtykker i at mitt barn deltar i studien.

Foreldres signatur:

Appendix 5 categorized comments

Teacher comment		Focus	Specificity	Mode
Title: C1 Your title addresses one side of the topic, can you rephrase it so it includes the two other aspects as well (in what ways is it a Lingua Franca and if it is a threat to other languages). You could simply say: "English – a world language".	C1	Local	Specific	Closed question/ advice
Introduction: C2 Interesting intro. It makes me want to read more 😊 C3 Can you somehow bring up the issue of threatening other languages in the intro as well? C4 Behind C5 better word would be 'country' 😊 C6 This is a good sentence to end an introduction with. You make it known to the reader that both English language and culture is still seen in countries around the world, and the reader now expects to read more about this. Good! C7 Take a second look at the task and see whether your introduction addresses all three questions you are expected to discuss. Check also whether you introduce all aspects you <u>actually discuss</u> in the text. Do you focus on historical reasons only in the main part?	C2 C3 C4 C5 C6 C7	Global/structure Global/structure Local Local Local	Vague Medium Medium Specific Specific Medium	Praise Open question Correction Advice Praise and explanation Advice and open question
1st paragraph: C8 Which enemies? C9 C11 to grow – grew – has grown (choose the correct past tense here). C10 Industrial C11 Handle means to take care of. I think you might be looking for the word 'trade', which is 'handel' in Norwegian	C8 C9 C10 C11	Local Local Local Local	Medium Medium Specific Specific	Closed question Instruction Correction Instruction, advice

C12 This paragraph nicely sums up the history of how GB became powerful. However, could you add a sentence at the end of the paragraph to make the transition to the next paragraph a little smoother? C13 Switch these paragraphs, and between them, add a bit about what happened in the 20 th century.	C12 C13	Global Global	Medium Medium/specific	Praise and open question Imperative
2nd paragraph C14 Could you use a linking strategy to start the paragraph with? C15 There is one reason, but there are many reasons 😊 C16 at C17 Rephrase this sentence C18 though C19 I am/ you are /he she it is/ we are /you are / they are 😊 C20 You present both the number of native speakers and the presence of English online as reasons for why English is a lingua franca. I agree! You do state that there are many reasons, but you only list two. Can you think of and refer to more reasons? Also, good use of sentence connectors, like 'the first reason is' C21 person C22 Try to cover other uses of English than just the internet.	C14 C15 C16 C17 C18 C19 C20 C21 C22	Global Local Local Local Local Local Global/Local Local Global	Medium Vague/Medium Specific Vague Specific Medium/specific Medium/specific Specific Specific	Open question/advice Instruction/correction Correction Imperative Correction Instruction/correction Explanatory/Advice/ Open question Correction Advice
3rd paragraph C23 Excellent, here you point at the negative sides to English becoming a global language. Could you start the paragraph with a sentence that makes the transition from the last paragraph better? C24 word choice	C23 C24	Global Local	Medium Vague	Explanatory/Open question

Appendix 6 categorized comments

C25 ?? Which place?	C25	Local	Specific	Instruction/Advice Closed question
4th Paragraph C26 Rewrite C27 missing word C28 I understand why you have included this. You want to illustrate that other languages also have made an impact on the English language. However, you need to rewrite this paragraph to make this clearer. I suggest a sentence or two at the beginning of the paragraph where you explain that other languages have influenced English, not only the other way around 😊	C26 C27 C28	Local Local Global	Vague Medium Medium/specific	Imperative Advice/instruction Imperative/Advice
Conclusion C29 I suggest that you rewrite this conclusion, where you sum up in short what you have discussed in your text. Perhaps start with something on the British empire? 😊	C29	Global	Medium/specific	Advice/ Closed question
Final comments Some things you do well: C30 Your essay is nicely organized in paragraphs – especially the introduction and most of the main part. (See advice on how to write a conclusion in the section on how to improve below.) C31 You answer important aspects of the question by pointing at some reasons to how English has become the most important language in the world, and you also point at some ways in which English is used as a lingua franca. C32 You use some formal linking words and sentence connectors such as “For example” and “On the other hand”, and these help you organize and connect your thoughts and ideas in a good manner. C33 You provide some examples to give ground to your opinions.	C30 C31 C32 C33	Global Global Local/global	Specific Specific Specific	Explanatory/praise Explanatory/praise Explanatory

Some things to work on in your final draft: C34 Work on how to write a good conclusion, because this is missing in your text. Remember that all new information and arguments should be addressed in your introduction and your main part, and that the purpose of a conclusion is to sum up the main elements as addressed in your text, without adding new information. C35 Answer all parts of the question. You forget to discuss whether English is a threat to other languages and cultures, and it can be argued that your paragraph about France and Japan may be a bit irrelevant. C36 Work on subject-verb agreement. (Is/are, has/have, etc.) Your text contains quite a few of these mistakes, and that makes your overall communication less precise. C37 Work on when and how to use the article “the”. Sometimes you use it when it should be left out, and other times it is missing in your sentences. C38 Pay more attention to the differences between English and Norwegian, and look up words in a dictionary if you are unsure of how to use them. The word “handle”, for example, does not mean the same thing in English and Norwegian. When using this word in its Norwegian sense it does not communicate clearly in English, and it may be difficult for the reader to understand what the sentence is really about.	C33 C34 C35 C36 C37 C38	Global Global Global Local Local Local	Medium/specific Specific Specific Specific Specific Medium	Explanatory Imperative/Instruction Imperative/ Instruction Imperative/ Instruction Imperative Advice
---	--	---	---	---

Appendix 7 Interview guide

Interview guide

Warm-up questions

1. What are your thoughts on writing in the English subject?
2. What types of texts do you prefer to write?
3. What do you find to be challenging when writing a text in English?
4. What strategies do you use to revise your own text before you hand it in?

Which types of feedback do students receive on drafts/texts?

1. What types of feedback other than grades are used when you write texts?

- peer assessment
- group assessment
- self-assessment
- Teacher-student conference
- written feedback

2. Can you describe how your teacher gives you feedback on your texts written in English?
3. In your experience, how often do you receive feedback on texts you have written?
4. How easy/difficult is it to understand the feedback you receive from your teacher? What is easy/difficult?
5. How do you understand what to do with the feedback?
6. In your experience, how often are you given the opportunity to work with the feedback to improve your text before it is given a final grade?
7. What do you do with the feedback after the grade is given, do you use it for subsequent writing? Or never looked at again?

Which types of feedback are most useful and why?

1. At this stage in the focus group we will look at different ways of providing written feedback. Which method of providing feedback do you prefer? (Margin comments, in-text comments, end comment, separate piece of paper).
2. The students are now given time to read the text, after which we will look at each feedback point, and ask about usefulness and understanding. Is the language clear, explicit and easy to follow up on?

Concluding questions

1. Aside from feedback from your teacher, what do you think can help become a better writer?
2. What type of feedback motivates you to continue working on a text? Your "ideal feedback", so to speak.
3. In your opinion, when is the best time to receive feedback? During the process of writing, or after the text is finished?