Assessing Oral Proficiency Levels of Second-Year Students of English at Radboud University

Michelle Everard

English Language and Culture

Second Semester, 2014-2015

BA Thesis Linguistics (LET-ETCB341)

Supervisor: Dr Rina de Vries

Second Reader: Dr Pieter de Haan

15 August 2015

Abstract

Oral proficiency can be tested in various ways. Students of English at Radboud University are assessed in pairs during a discussion and are assessed with the CEFR. Although the CEFR is useful and effective as an assessment tool, it is still not clear where the differences in oral proficiency levels of students lie. This research deals with the following research question: How different are the oral proficiency levels of second-year students of English at Radboud University? The hypothesis is that students who are more fluent are expected to make fewer mistakes, demonstrate a wider vocabulary range, and are also expected to be less hesitant. In order to answer this research question, four CEFR scales are used as a starting point for in-depth analysis of the students' oral proficiency levels. Each scale is linked to certain features like number of hesitations, and lexical errors, in order to complement the descriptors and assessment. By using both the CEFR and the features, the differences between the oral proficiency levels have become clear. The results have proven that the hypothesis is incorrect.

Keywords: Oral proficiency, CEFR, Assessment, second-year students of English, Radboud University, fluency, EFL.

Table of Contents

Introduction

There are approximately 427 million L1 speakers and 950 L2 speakers of English, which makes it one of the most spoken languages in the world (Saville-Troike, 2012, p. 9). English is often taught at schools and universities as an L2 language. It is also taught and tested at Radboud University Nijmegen where multiple aspects of the English language are assessed. Oral proficiency, for instance, is assessed in the first and second year. They are tested in fluency exams, in which two students discuss two newspaper articles with each other for approximately twenty minutes. Two examiners assess them with the help of their notes made during discussion and the CEFR, a framework that has gained a lot of ground over the years despite a lot of criticism voiced by many scholars. The students are assessed on several CEFR scales such as Vocabulary Range, Vocabulary Control, and Turn Taking. In these exams, it became clear that there are differences in the students' oral proficiency levels. Some students scored B2s while others were assessed with C1 or even C2. It can thus be said that, generally, there is a difference in oral proficiency levels in terms of CEFR levels, though this statement says nothing concrete about where the differences lie.

In this research, I will complement the CEFR assessments with in-depth analysis in order to answer the following research question: How different are the oral proficiency levels of second-year students of English at Radboud University? The hypothesis of this thesis is that there will be a clear distinction between advanced and less advanced students. The students who are more fluent are expected to make fewer errors and restarts, to demonstrate a more varied use of vocabulary, and are also expected to be less hesitant than the students who are less fluent.

This thesis is structured as follows: the first will cover the Common European Framework of Reference for Languages. It will explain how it has been set up, what scales make up the Framework, and will cover the criticism it has received over the years. The second chapter is about Proficiency. In this chapter, definitions and different assessments are discussed. The third chapter deals with my research and results. In this chapter, the methodology, participants, materials, and the results per scale will be discussed. In the fourth chapter, the results are analysed and discussed in order to answer the research question. This thesis will be rounded off with a conclusion, followed by the appendices.

The Common European Framework of Reference for Languages

The Common European Framework of Reference for Languages (CEFR) is a framework devised by the Council of Europe in 2001. The Framework has been set up with attention to following three criteria. Firstly, it should be a comprehensive framework, which means that it covers all skills and uses of language possible (Council of Europe, 2001, p. 7). The Framework must also be transparent, meaning that it is explicit and clear, and "readily comprehensible to users" (p. 7). Finally, it must be coherent and should not contradict itself (p. 7).

Next to these criteria, the Framework has a few aims. One is to encourage practitioners of every different kind to ponder about questions involving (second) language acquisition, for instance "how does language learning take place?" and "what can we do to help ourselves and other people to learn a language better?" (p. i). The other main aim is to make it easier for both teachers and learners to set clear goals and, in turn, create ways to achieve those goals. Ultimately, the Framework wants to inspire users and certainly not impose particular language strategies or determine how a user should use the Framework. "We are raising questions, not answering them," sums up the Framework's attitude (p. i).

Setting up a Framework

From 1970s onward, the Council began setting up the Framework. It started with the threshold-levels and was later expanded to the Framework that we are familiar with today. The Framework was not based on empirical evidence, i.e., L2 learner data because there was no corpus data available to base the Framework on at the time (Council of Europe, 2001, p. 21). The devisors of the Framework thus had to rely on other sources and turned to teachers' perceptions, which was considered the next best thing.

The first stage of devising the Framework was called "Intuitive Analysis" (North, 2000, as cited in Fulcher, 2004). At this stage, existing scales from, for instance, International English Testing Service (IELTS) and Association of Language Testers in Europe (ALTE), were put together and gaps were filled with new scales (North, 2000, p. 181). Then the Council of Europe had teachers evaluate those scales. The scales that were evaluated most consistently by teachers were ultimately compiled into questionnaires. The next stage was qualitative analysis with questionnaires in which teachers were asked to put certain skills to their most fitting levels. Skills that were deemed unfitting were removed from the pool. The final stage entailed replication of the former tests, which resulted in a correlation of .99 (Fulcher, 2004, p. 257). After the results had been thoroughly studied, quality-control of descriptors took place in which problematic scales were further analysed and ultimately

brought down to 212 descriptors (North, 2000, p. 260, 271). The descriptors underwent even more scrutiny in the next stage of setting up the Framework. The descriptors were judged on whether they were coherent and "whether progress up the scale in each category was logical" (p. 271). Consequently, the teachers that were involved were asked again to test the descriptors, but were now asked to assess their students. It was concluded from this study, though admittedly on a "very thin basis", that teachers could find a 'cut-off point' at which a learner was not able to do a certain task (p. 333). Clearly, it took a lot of research, time, and effort in order to put the Framework together.

Many, however, remain sceptical of the Framework's validity and basis as the primary concern for many scholars is the Framework's lack of empirical basis. Alderson (2007), Fulcher (2004), Hulstijn (2007), Little (2006, 2007), and North (2007, 2014a, 2014b) have pointed out this gap. Fulcher (2004) argues that rising importance of CEFR might cause teachers to be under the impression that the CEFR is built on empirical evidence rather than agreed perceptions. 'Common' in the acronym CEFR thus refers to a common agreed perception and is far from a natural progress by actual learners (Fulcher, 2004). Alderson (2007) also criticises the Framework's basis and lack of empirical research as it "is giving rise to increasing misgivings about the applicability of the Framework in its current form" (p. 660). While North (2007) acknowledges that the scales are not based on L2 learner data but rather on agreed perceptions, he still disagrees with the aforementioned scholars. North argues that the scales do have "a good degree validity" because Swedish learners of Finnish self-assessment was similar to the CEFR scales (p. 657). Whether this one instance can be evidence in favour of validity of the scales, is still up for debate, but North does not regard the lack of L2 learner data as a reason to reject the Framework in its entirety.

In order to remedy the gap, many call for more research concerning the Framework, but also L2 learner data. Hulstijn (2007) encourages the use of L2 learner data and corpus research in order to strengthen the Framework, as "it is high time that researchers of SLA, research of language assessment, and corpus linguists paid attention to each other's work and engaged in collaborative research" (p. 666). While North (2007) supports the idea of putting the Framework on the research agenda, he remains sceptical that a grand scale empirical framework research project will never happen as SLA research is not concerned with such a topic at this point in time.

Why do we need the CEFR?

In explaining why the CEFR is needed, the Council of Europe refers to a speech made at the Intergovernmental Symposium of 1991 at Rüschlikon. According to the Council, more attention should be given to language learning and teaching so that international communication, respect, and working relations may be improved. In order to achieve these goals, language learning must be encouraged from primary school to adult education. The Framework might also help "facilitate co-operation among educational institutions of different countries […], provide a sound basis for the mutual recognition of language qualifications and to situate and co-ordinate efforts" of learners and teachers (Council of Europe, 2001, p. 5).

These early wishes for the Framework have been incorporated in aims of the final product, which is apparent in the Council's aim concerning syllabi. They hope that the Framework will influence syllabuses and courses to the extent that they will be more transparent and geared towards "international co-operation in the field of modern languages" (p. 1). Clearly it is desired that the Framework will be much more than just a tool for language learners and teachers. Some, however, are wary of the Framework and its political intentions. Especially Fulcher (2004) is afraid that the Framework's influence is stretching too far and claims that "there is a strong political agenda at work" (p. 262). Fulcher is not convinced that EU member states are keen on harmonisation and argues that harmonisation will jeopardise diversity (p. 254, 255). Fulcher is wary of the Framework's lasting impact and worries that the Framework's increasing status as an important language tool will leave little room for critics to reject the Framework and criticise its many shortcomings (p. 260). Similarly, Davidson and Fulcher (2007) express their concerns that the Framework is well on its way of becoming 'the' framework, despite the Council's vehement disclaimer that it should not be regarded as such.

The Scales

The CEFR has multiple scales that can be used for (self-)assessment. These illustrative descriptors are divided into different competences, strategies, domains, and activities, which is the so-called "horizontal mapping" of the Framework (Council of Europe, 2001, p. 16). The vertical dimension, however, is the aspect that the Framework is most widely known for. Each scale has a minimum of six levels, from A1, being the lowest, to C2, the highest level on the scale. Some scales may have +-levels that further define certain competencies. Both dimensions can be used separately or in combination, but are equally as important within the Framework.

Language learners are categorised in several ways in the Framework. A learner that scores the lowest on the Framework is commonly known as a basic user, whose sublevels are divided into A1: Breakthrough and A2: Waystage. B-level students are called independent users, with B1: Threshold and B2: Vantage as their sublevels. Lastly, C-level students are called proficient users, with C1: Effective Operational Proficiency and C2: Mastery (Council of Europe, 2001, p. 23). Each user has its own distinct features that belong to a specific level of a illustrative descriptor. It is assumed that when a learner has reached, for instance, B1 on a illustrative descriptor, the learner is capable of what is described in A1 and A2, which suggests that within a particular scale, progress is linear. This view, also known as the language ladder, is often challenged. The "ice-cream cone" image is said to be a more suitable representation, especially with the CEFR's horizontal and vertical dimensions kept in mind (North, 2014a, p. 101). Similarly, Hulstijn (2007) rejects linear language learning as it is presented in the Framework, especially because there is no empirical evidence that progress is made in that manner (p. 666).

An Example of a CEFR Scale: Overall Oral Production

	OVERALL ORAL PRODUCTION
C2	Can produce clear, smoothly flowing well-structured speech with an effective logical structure which helps the recipient to notice and remember significant points.
C1	Can give clear, detailed descriptions and presentations on complex subjects, integrating sub-themes, developing particular points and rounding off with an appropriate conclusion.
B2	Can give clear, systematically developed descriptions and presentations, with appropriate highlighting of significant points, and relevant supporting detail.
B2	Can give clear, detailed descriptions and presentations on a wide range of subjects related to his/her field of interest, expanding and supporting ideas with subsidiary points and relevant examples.
B1	Can reasonably fluently sustain a straightforward description of one of a variety of subjects within his/her field of interest, presenting it as a linear sequence of points.
A2	Can give a simple description or presentation of people, living or working conditions, daily routines, likes/dislikes, etc. as a short series of simple phrases and sentences linked into a list.
A1	Can produce simple mainly isolated phrases about people and places.

(Council of Europe, 2001, p. 58)

The 56 illustrative scales are based on several skills, namely, reading, writing, listening, and speaking. These categories are then further divided into two categories, language use and strategy scales and competency scales. Examples of language use and strategy scales are 'Overall Oral Production', 'Goal-Oriented Co-Operation', and 'Reading

for Information and Argument'. 'Turntaking', 'Thematic Development', and 'Grammatical Accuracy' are competency-based scales. Each illustrative scale is intended to be as context-free as possible in order to be applicable for as many users as possible. The intention to remain context-free has several consequences for the Framework as it leads to an incomplete framework.

The Framework does not cover electronic communication like texting and the internet (Little, 2011, p. 386). The Framework also does not take abbreviated language into account even though it is used in everyday life (p. 386). Additionally, there are still gaps in the scales, often at A1 and A2 level or C1 and C2 level, like in the Sustained Monologue: Putting a case (e.g. in a debate) (p. 59). Monitoring and Repair also has no descriptors available for A1 and A2. Information Exchange has no available descriptors for C1 and C2 level and at those levels refers to the B2 descriptor instead. The Council readily acknowledges these gaps, and suggests that users fill in the gaps for themselves, though also warns that some gaps will never be filled or remedied as the scale works fine without it (2001, p. 37).

Some users may find problems in the non-specific approach of the Framework, while others might be delighted with this as it allows for a lot of freedom and flexibility. Fulcher (2004), Davidson and Fulcher (2007), and Alderson (2007) are opposed to this non-specific approach. Fulcher (2004) argues that the Framework is "so abstract that is not a framework, but a model" (p. 258). Davidson and Fulcher argue that next to being too abstract and vague, the scales are inconsistent (p. 234). Some descriptors are clear and refer to specific situations, while other scales vaguely hint at situations or do not refer to anything at all. Alderson (2007) additionally states that abstract descriptors are "couched in language that is not easy to understand" resulting in the fact that the Framework is not particularly user-friendly (p. 661), which was one of the criteria which the Framework was built on. North (2007) acknowledges that the Framework is often criticised for its shortcomings and alleged vague approach, though still defends the Framework. Firstly, North (2000) does think that the CEFR is user-friendly as it was comprised with help from teachers (p. 335). He also states that the scales are set up to be as specific as possible, but that that task is "a tall order" (2007, p. 658). A framework can only be as exhaustive and detailed in order to remain context-free and accessible for everyone, underlines North.

North clearly is in line with what the Council had in mind for the users of the Framework, namely that the user himself should actively use the framework. They might add or leave out certain parts of the framework to make sure that it works best for them, their context, and their aims. The Framework is not intended to be used without any consideration

on the user's part. North (2014b) states "the CEFR is a heuristic tool, but it is not the answer to all problems" (p. 243). It is an "inspiration, not a panacea […]" and should be "critically consulted" (p. 245). While the other scholars seem to be set on rejecting the Framework because of its non-specific approach, North underlines the original intention of the creators of the Framework.

There are also other concerns about the descriptors. Like Alderson (2007), Little (2006) suggests that the framework is not user-friendly, because some scales are unsuitable for younger learners. The descriptors used in higher levels are only suitable for older learners due to the skills described in higher levels. Little (2011) thus calls for a revision of the higher levels so that they can be used by younger learners as well. Little (2006) also criticises the validity of the benchmarks and the scales itself (p. 186). He is opposed to fluency being scaled in how hesitant a speaker is, since many native speakers can be hesitant speakers as well despite being perfectly fluent (p. 186). Scales like these need more research and may need expansion, just like the rest of the Framework, according to Little. He (2007) also notes that the Framework is not suitable as a basis for tests, which Alderson et al. (2006) also found with their Dutch CEFR Construct Project. Moreover, Little claims that the CEFR can only function as a "starting point" for test designs as it is not language specific (p.649). Alderson et al (2006) argue that there is not "sufficient guidance" to base tests on CEFR levels, though they state that it is a "tentative conclusion" as more research is needed concerning CEFR and testing (p. 21).

The Future

While North (2007) expressed his doubts about research regarding the CEFR, multiple projects have been set up to further research CEFR and its gaps. SLATE (Second Language Acquisition and Testing in Europe) focuses on answering this research question: "which linguistic features or learner performance (for a given target language) are typical at each of the six CEF levels?" with the help of learner corpora of multiple languages (Edmonds & Leclercq, 2014, p. 15). Like SLATE, the English Profile Programme (EPP) also tries to answer that question and has already found evidence of "criteria features of syntactical, morphological, and lexical use that distinguish between levels" (Hawkins & Filipoviçm, 2012; Salamoura & Saville, 2011, as cited in North, 2014a, p. 24). Thus far, both of these studies have given support to the CEFR scales and their validity (p. 24).

In summary, the Common European Framework of Reference for Languages is heavily debated phenomenon. While the Council's aims and objectives concerning the Framework are abundantly clear, attitudes towards the Framework remain mixed. Many scholars criticise the descriptors, the impact of the Framework, and lack of empirical foundation which in turn leads them to question the validity of the Framework. Even though the Council of Europe's intention is to encourage discussions on language learning and for the Framework to be accessible to as many users as possible, scholars reject the Framework because of its shortcomings and claim that it is all but user-friendly and accessible. Thus scholars urge for more research, especially into L2 learner data and how that data can be linked to a framework like CEFR. Though, recent research into the CEFR has given some support to the Framework and that hopefully will counter many scholars' criticism and will usher in a time where the Framework will be celebrated instead of criticised.

Proficiency

Defining Proficiency

Around the 1950s, linguists thought grammatical competence only to be an indicator of proficiency (as cited in Edmonds & Leclercq, 2014, p. 6). Hymes (1972) insisted that communicative use is equally important as grammatical competence (as cited in Edmonds & Leclercq, 2014, p. 6). Lado's (1957) expanded the definition of proficiency even further. His definition is made up of four language 'elements', pronunciation, grammatical structure, lexicon and cultural meaning, and four language 'skills', namely speaking, listening, writing and reading (as cited in Young & He, 1998, p. 4). Canale and Swain (1980) further defined language proficiency as they added linguistic, pragmatic, discourse, and strategic competence as components of proficiency (as cited in Young & He, p. 4). Thomas' (1994) definition of proficiency is "a person's overall competence and ability to perform in L2" (as cited in Zhang, 2015, p. 79). Thomas' definition is similar to Briere (1972). Briere states that language proficiency is "the degree of competence or the capability in a given language demonstrated by an individual at a given point in time independent of specific textbook, chapter in book, or pedagogical method" (as cited in Esteki, 2014, p. 1522).

Skehan (1989) was first to define language proficiency with three core components complexity, accuracy, and fluency, also known as CAF (as cited in Edmonds & Leclercq, p. 8). Complexity is how "varied" and "elaborate" a speaker's language is during a specific task (Ellis, 2003, as cited in Edmonds and Leclercq, p. 8). Accuracy naturally means correct use of language without any errors. Fluency, however, is the most challenged component as its definition is not straightforward. Fluency can be defined as how closely learners' speech resembles a native speaker or how hesitant a speaker is (Lennon, 1990; Ellis, 2003, as cited in Edmonds and Leclercq, p. 8).

Hulstijn (2011) elaborately defines language proficiency as the following:

"the extent to which an individual possesses the linguistic cognition necessary to function in a given communicative situation, in a given modality (listening, speaking, reading or writing). Linguistic cognition is the combination of the representation of linguistic information (knowledge of form-meaning mappings) and the ease with which linguistic information can be processed (skill). Form-meaning mappings pertain to both the literal and pragmatic meanings of forms in

decontextualised and socially-situated language use, respectively)" (as cited in Edmonds & Leclercq, 2014, p. 7).

He argues that language proficiency is made up of peripheral and core components. Peripheral components entail, for instance, strategic competence and core components are made more up of linguistic cognition, which he further separates into two concepts, basic language cognition (BLC) and higher language cognition (HLC) (as cited in Edmonds & Leclercq, 2014, p. 7). Basic language cognition is made up of frequent lexical items and common grammatical structures. It is also implicit knowledge that most adult L1-speakers are familiar with (as cited in Edmonds & Leclercq, 2014, p. 7). Higher language cognition serves as a "complement" to BLC. Unlike BLC, HLC is comprised of less frequent lexical items and structures and are thus more complex than BLC elements (as cited in Edmonds & Leclercq, 2014, p. 7). Edmonds and Leclercq are of the opinion that Hulstijn's definition is the most accurate in describing L2 proficiency.

Zhang (2015) and Esteki (2014) add another component to the definition of proficiency. Both argue that implicit and explicit knowledge are part of language proficiency. Implicit learning is done without awareness, while explicit learning is done with awareness and is "product of language learning" (Ellis et al., 2009, as cited in Esteki, 2014, p. 1520). Not all scholars agree that there is such a phenomenon at work. Schmidt (2011) does not believe that implicit knowledge exists because "people learn about things they attend to and don't learn much about things that they don't attend to" (as cited in Esteki, 2014, p. 1520-21).

Zhang (2015) discusses Han and Ellis' (1998) research into explicit and implicit knowledge . Their research has given some support to Bialystok's (1982) claim that language proficiency is made up of explicit and implicit knowledge, i.e., "unanalysed and analysed knowledge" (as cited in Zhang, 2015, p. 80). Zhang notes, however, that even though some studies support this idea, a general conclusion is lacking and this theory still needs further investigation. Esteki (2014), however, does assume that implicit knowledge plays a big role in language proficiency, though, like Zhang, argues that research into the relation between explicit knowledge and proficiency is needed.

Different Proficiency Assessments

Proficiency can be assessed in different ways. Weir (1990) makes a distinction between discrete language testing and global assessment which is also known as integrative

assessment. Discrete language testing breaks language apart and assesses its components in isolation. In global assessment, however, the focus is on the overall performance of the learner. Hulstijn (2010) defines global assessment as an assessment that tests a "mixture of knowledge and abilities" (as cited in Edmonds & Leclercq, 2014, p. 12). According to Oller (1979), though, discrete language testing falls short because the assessment is based solely in isolation without any regard to "a larger context of communication", which integrative assessments do pay attention to (as cited in Weir, 1990, p. 2). Oller thus favours integrative assessments over discrete language assessment.

Learners can also be tested individually or in groups or pairs. He and Young (1998) are in favour of language proficiency interviews (LPIs), which are done with one learner and one assessor, preferably a native speaker. These interviews are largely question-and-answer based. They argue that having a native speaker assess a learner is the best way to judge whether someone is proficient or not (p.1).

Assessing students in pairs has multiple advantages. There is evidence that working in pairs motivates students engage more in conversation and thus improve their oral proficiency (Taylor, 2000, as cited in Davis, 2009, p. 369). It also does not feature the restricting question-and-answer style used in interview assessment, which means that pair assessment is closer to real conversation (Egyud & Glover, 2001; Johnson, 2001; van Lier, 1989; Young & Milanovic, 1992, as cited in Davis, 2009, p. 369). Consequently, pair assessment also leaves more room for variation as there is hardly any restriction (Skehan, 2001; French, 2003, cited in Davis, 2009, p. 369). Moreover, working in pairs is not uncommon is class room settings, which makes this format suitable for assessment (Davis, 2009, p. 369).

Edmonds and Leclercq (2014) argue that the choice between individual or group testing depends on practical matters. The educative system does not permit individual assessments as they are often not enough resources for teachers to assess learners individually (p. 13). Furthermore, researchers often cannot find the time to assess participants individually (p. 13).

What should be tested?

What is tested in proficiency assessments depends on what the assessor thinks is the goal a learner should strive for. He and Young (1998) are of the opinion that native speakers are the best to judge how proficient a learner is. They thus adhere to the concept of the 'ideal native speaker'. According to them this ideal must function as a yardstick for learners. This is a challenged view. The Council of Europe states that the ideal native speaker is "utopian"

because language learning is a never-ending process as no learner can achieve mastery in all language skills (2001, p. 169). Like the Council of Europe, Ross (1992) questions the reliability of a native speaker's intuition (p. 174).

Weir (1990) argues that a shift has occurred in testing. Formerly, learners were tested on their linguistic accuracy, though now the focus is more on their communicative skills in particular contexts (p. 9). In other words, learners are assessed on their demonstration of their skills in their own right and no 'native speaker ideal' is involved.

Problems with Assessment and Tests

Bachman and Palmer (1996) argue that reliability, construct validity, authenticity, and interactiveness are the four necessary components of a "useful language test" (as cited in He & Young, 1998, p. 1). Reliability is a big problem within assessment, claim Bachman and Palmer. Interviewers can disagree on a learner's results, which endangers the reliability of a test. Disagreement can be circumvented by using rating scales, Bachman and Palmer argue (as cited in He & Young, 1998, p. 2). Valid tests can be made when there is an understanding on what it exactly tests. For instance, in order to test oral proficiency accurately, a researcher needs to have a clear idea of what oral proficiency entails (as cited in He & Young, 1998, p.2). Unlike Edmonds and Leclercq (2014), who think validity, reliability, and practicality are important in assessments, Bachman and Palmer (1996) add to other necessary components to tests, namely authenticity and interactiveness. Authenticity is an important factor because it can lead to generalisations on proficiency (as cited in He & Young, 1998, p. 2). Interactiveness is not about how the participant and assessor interact, but the way in which the participant "draws on different kinds of knowledge" which are, for instance, "knowledge of a second language, knowledge of how to overcome communication difficulties in performance (strategic competence) [and] knowledge of how to organize and plan a task (metacognitive strategies)" (as cited in He & Young, 1998, p. 3). All of these criteria are important in assessments, though there are still other problems concerning testing.

Many scholars are concerned with how assessments reflect real life conversation. Weir (1990) argues that it is nearly impossible to recreate "real-life communication" in test environments, which makes it difficult to make reliable and valid demands of a learner in such a setting (p. 16). Bachman (2002) argues in the same vein that assessments do not test what is taught in class (as cited in North, 2014b, p. 159). While learners deal with real life-tasks in class, tests are not geared towards testing those exact tasks. Lantoff and Frawley (1985, 1988) also argue that learning criteria used to define assessments "are not anchored to

any set of features evolving from natural communication" (as cited in Ross, 1992, p. 174). Moreover, they argue that assessments not only assess proficiency, but how well a learner can get through a test.

Additionally, Weir (1990) argues that researchers and teachers need to be wary of making all-conclusive statements about "similar communication tasks" based on specific tests (p. 17). Even 'similar' tasks can be different in the way learners deal with them, so all-conclusive statements cannot be easily made, according to Weir. Similarly, learner performance during assessments may vary depending on the task, time, interlocutors, and environment (Davis, 2009, p. 368). Examiners consequently cannot make generalising conclusions about learners and their performance.

Common European Framework of Reference for Languages as a Testing Tool

Proficiency rating scales like the Common European Framework of Reference for Languages can also be used to assess proficiency. The CEFR can be used for assessment in three ways. It can be used for identifying what needs to be tested, how learner performance can be interpreted, and how comparisons can be made (Council of Europe, 2001, p. 178). The Council argues that the Framework "seeks to provide a point of reference, not a practical assessment tool" (p. 178). North (2014a) states that users can use the CEFR as a starting point for assessment, but should definitely think about what they want to get out of the assessment and what needs to be tested.

Osborne (2014) is critical of CEFR as a tool to measure proficiency. He argues that parts of some descriptors are "somewhat random" as the distinction between "very noticeable" and "very evident" might be confusing to users of the Framework (p. 57). He set up a study to research whether assessors use the Framework in the same understanding and if these descriptors caused problems for users. His research found that assessors used the Framework with "relatively strong agreement" and were able to "reliably find the cut-off point at which learners are not able to the things described" (p. 62). Users of the Framework thus search for the point a learner falls short, which is the exact opposite of what the Framework wants to encourage, namely staying positive, says Osborne (p. 62). Similarly, North (2000) also gives support to the Framework's reliability (as cited in North, 2014a, p. 213). His research has found that 73.5% of the teachers' assessments corresponded with each other.

Defining the term 'proficiency' is challenging. Scholars define proficiency either narrowly or very broadly. Thomas (1994), for instance, describes proficiency as "a person's overall competence and ability to perform in L2" (as cited in Zhang, 2015, p. 79), while

Hulstijn (2011) adopts a very elaborate definition of proficiency as he adds multiple dimensions to his definition. Ultimately, there is no 'official' way of explaining what proficiency is as scholars do not agree on what proficiency entails. It is also hard and nearly impossible for scholars to make all-conclusive statements about proficiency because the term is not set in stone. Another problem is that too many factors influence assessments and learner's performance. Proficiency can be tested in multiple ways, and again, there is no 'official' or 'right' way to assess a learner and every examiner has their own preferences. There are several kinds of ways to assess proficiency. Two examples are discrete language testing and integrative assessment. Proficiency scales like the CEFR might be used as well. While some have expressed their concerns whether the CEFR can be used appropriately and accurately, the Framework has been credited some reliability by multiple researches. The creators of the Framework, though, mostly encourage users to adapt the Framework to their own needs instead of using it as a "practical language tool" (Council of Europe, 2001, p. 178).

Research and Results

The research question for this thesis was: how different are the oral proficiency levels of second-year students of English at Radboud University? The hypothesis was that there would be a clear distinction between advanced and less advanced students. The students who were more fluent were expected to make fewer errors and restarts, to demonstrate a more varied use of vocabulary, and were also expected to be less hesitant than the students who were less fluent. In order to confirm this hypothesis, the following research was set-up and carried out.

Participants

Eight second-year students of English Language and Culture at Radboud University were chosen to take part in this research. First-years were not chosen for this research because those had not been students for a long time at the start of this project. They also have not done a fluency exam at that point in their studies. Third-year students were also unsuitable for this project due to the fact that some students had been abroad while many students had not, which makes it a far less homogenous group than second-year students. Second-year students have not been abroad yet and have roughly spent the same amount of time improving and working on their fluency in, for instance, the Oral Communication Skills courses. They also should have a certain level of proficiency after studying English for one and a half year otherwise they would not have made it to their second year. Second-year students should generally have a C1-level according to the Oral Communication Skills teachers.

Even though they might have been in the same classes, some students improve at a faster rate than other students. In order to test my hypothesis, four relatively less fluent students and four relatively more fluent students were chosen, though they were not notified of the selection criteria. The choice of these eight students was based on their fluency grades and their Oral Communication Skills teachers' perceptions. The students were paired up according their level. This was done in cooperation with Dr de Vries and with the help of their Fluency grades. Jane and Lena scored 8 and 6, John and Marie both scored a 4, Anna and Celia scored 7 and 4.5, Dawn and Liz scored 7.5 and 7. They were not aware of the fact that they were paired up with a student of a comparable level. The students were aware that my research involved fluency and the CEFR which they were already familiar with. The students' names were changed throughout this research in order to protect their privacy.

Materials

 The students discussed "Tories to announce resits for pupils who fail end of primary school exams" from the Guardian (Appendix A) for approximately ten minutes in pairs. This article was chosen because it seemed long enough for a ten-minute discussion. Moreover, the article showed two sides to the argument, which meant that it was objective and students could still take a side. The article was also suitable because of its topic, namely testing, could easily be applied to university or secondary school as well.

Methodology

 Prior to the discussion, the students received instructions on a sheet and were asked to read it thoroughly (see Appendix F). On the same sheet they were asked to give permission for my recording them and using their data for my thesis. Then the students had the opportunity to carefully read the Guardian article. They were allowed to make notes. After they had finished reading the article, they had the opportunity to ask questions about things they did not understand. Finally, I repeated the instructions on the permission sheet and then instructed to start their discussion. The entire discussion was filmed and recorded with permission of the participants gained prior to the experiment.

 The discussions were done under my supervision, so that the students could ask for help if the conversation stalled, but otherwise I tried to intervene as little as possible. If the conversation stalled, I had some questions prepared so I could help them get back into the discussion. The same questions were used during each discussion and tried to ensure that each pair discussed the same topics.

 After the discussions had taken place, all discussions were transcribed orthographically. The transcriptions can be found in the Appendix B-E. I further analysed the transcribed data using the programmes CorpusTools (O'Donnell, 2007) and AntConc (Anthony, 2014).

 Several CEFR scales were chosen as a starting point for this research. The scales Vocabulary Range, Vocabulary Control, Spoken Fluency, and Propositional Precision were selected. The reason for picking these four scales was that the first three scales were also used during normal Fluency exams. The scale Propositional Precision was picked because it covers more advanced language skills compared to the other scales. It was fitting because the scale is about expressing opinion and my research is based on discussions between students.

 For each scale a number of features related to the CEFR scales were chosen that were further investigated. For Spoken Fluency I looked at restarts, hesitations, numbers of

sentences and words spoken, and number of turns. For Vocabulary Control, I looked at lexical errors. For Vocabulary Range, I looked at how much of the vocabulary from the article was used by the students, their Type-Token ratio, and their Lexical Density. Lastly, for Propositional Precision, I investigated how students voiced their opinions and which degree adverbs they used.

      After collecting all the data, each student was analysed per scale and compared them with the other students. Finally, every students was graded with the relevant CEFR scales and concluded my research by answering my research question.

General Observations

      Every participant stayed within the scope of the article except for Lena and Jane. All participants contributed to their discussion by asking questions, asking for clarification, and by bringing in new points.

Spoken Fluency

CEFR Spoken Fluency Scale

	SPOKEN FLUENCY
C2	Can express him/herself at length with a natural, effortless, unhesitating flow. Pauses only to reflect on precisely the right words to express his/her thoughts or to find an appropriate example or explanation.
C1	Can express him/herself fluently and spontaneously, almost effortlessly. Only a conceptually difficult subject can hinder a natural, smooth flow of language.
B2	Can communicate spontaneously, often showing remarkable fluency and ease of expression in even longer complex stretches of speech.
	Can produce stretches of language with a fairly even tempo; although he/she can be hesitant as he/she searches for patterns and expressions, there are few noticeably long pauses. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without imposing strain on either party.
B1	Can express him/herself with relative ease. Despite some problems with formulation resulting in pauses and 'cul-de-sacs', he/she is able to keep going effectively without help.
	Can keep going comprehensibly, even though pausing for grammatical and lexical planning and repair is very evident, especially in longer stretches of free production.
A2	Can make him/herself understood in short contributions, even though pauses, false starts and reformulation are very evident.
	Can construct phrases on familiar topics with sufficient ease to handle short exchanges, despite very noticeable hesitation and false starts.
A1	Can manage very short, isolated, mainly pre-packaged utterances, with much pausing to search for expressions, to articulate less familiar words, and to repair communication.

(Council of Europe, 2001, p. 129)

The scale Spoken Fluency is concerned with 'natural' speech flows, hesitations, and pauses. Students can be rated from A1 to C2 with additional A2+, B1+ and B2+ levels.

Table 1: Duration and Word Counts

Pairs	Duration of Discussion	Number of spoken words	Words per minute
John and Marie	10m 4s (604s)	1530	152
Dawn and Liz	11m 6s (666s)	1547	139
Jane and Lena	11m 40s (640s)	2116	198
Anna and Celia	11m 45s (645s)	1609	150

Table 2: Turns, Word Count, Spoken Sentences

Participants	Turns	Number of spoken words	Number of spoken sentences	Average x words per sentence
John	37	687	68	10.1
Marie	35	843	61	13.8
Dawn	33	803	61	13.1
Liz	34	744	61	12.1
Jane	37	1059	68	15.5
Lena	36	1057	82	12.8
Anna	43	1162	86	13.5
Celia	36	447	42	10.6

Table 3: Hesitations

Participants	"Erm"	"Er"	Combined hesitations
John	28	0	28
Marie	8	11	19
Dawn	16	14	30
Liz	26	5	31
Jane	14	34	48
Lena	3	5	8
Anna	2	0	2
Celia	0	0	0

Table 4: Restarts

Participant	Restarts (--)
John	9
Marie	17
Dawn	13
Liz	10
Jane	31
Lena	22
Anna	35
Celia	4

Vocabulary Range

CEFR Scale Vocabulary Range

	VOCABULARY RANGE
C2	Has a good command of a very broad lexical repertoire including idiomatic expressions and colloquialisms; shows awareness of connotative levels of meaning.
C1	Has a good command of a broad lexical repertoire allowing gaps to be readily overcome with circumlocutions; little obvious searching for expressions or avoidance strategies. Good command of idiomatic expressions and colloquialisms.
B2	Has a good range of vocabulary for matters connected to his/her field and most general topics. Can vary formulation to avoid frequent repetition, but lexical gaps can still cause hesitation and circumlocution.
B1	Has a sufficient vocabulary to express him/herself with some circumlocutions on most topics pertinent to his/her everyday life such as family, hobbies and interests, work, travel, and current events.
A2	Has sufficient vocabulary to conduct routine, everyday transactions involving familiar situations and topics.
	Has a sufficient vocabulary for the expression of basic communicative needs. Has a sufficient vocabulary for coping with simple survival needs.
A1	Has a basic vocabulary repertoire of isolated words and phrases related to particular concrete situations.

(Council of Europe, 2001, p. 112).

The Vocabulary Range descriptors involve how advanced a learner's vocabulary is. Idiomatic expressions, colloquialisms, lexical gaps, circumlocution, repetition, and connotative meaning are all part of this scale. Students can be rated from A1 to C2 with an additional A2+ level. In order to measure the students' vocabulary range, the students' use of vocabulary from the article was studied and their type-token ratio and lexical density were calculated.

The Type-token ratio is the number of unique words used (type) divided by the total number of words (tokens). The higher the type-token ratio, the wider a person's vocabulary range is. Lexical density is the percentage of lexical words used.

Table 5: Lexical Density

Participant	Lexical density
John	48.03%
Marie	42.34%
Dawn	43.96%
Liz	43.54%
Jane	43.05%
Lena	40.11%
Anna	39.58%
Celia	44.51%

Table 6: Type-Token Ratio

Participant	Type	Token	Type-Token Ratio
John	223	687	0.32
Marie	225	843	0.27
Dawn	242	803	0.30
Liz	199	744	0.27
Jane	289	1059	0.27
Lena	287	1057	0.27
Anna	303	1162	0.26
Celia	178	447	0.39

Table 7: Vocabulary Matches

Participant	Vocabulary matches with article	Unique words	Percentage of vocabulary matches
John	91	223	40.81%
Marie	96	225	42.68%
Dawn	115	242	47.52%
Liz	92	199	46.23%
Jane	99	289	34.26%
Lena	104	287	36.24%
Anna	112	303	36.96%
Celia	77	178	43.26%

Vocabulary Control

CEFR Scale Vocabulary Control

	VOCABULARY CONTROL
C2	Consistently correct and appropriate use of vocabulary.
C1	Occasional minor slips, but no significant vocabulary errors.
B2	Lexical accuracy is generally high, though some confusion and incorrect word choice does occur without hindering communication.
B1	Shows good control of elementary vocabulary but major errors still occur when expressing more complex thoughts or handling unfamiliar topics and situations.
A2	Can control a narrow repertoire dealing with concrete everyday needs.
A1	No descriptor available

(Council of Europe, 2001, p. 112).

The scale Vocabulary Control is concerned with correct use of vocabulary. Students can be rated from A2 to C2, but the A1 descriptor is not available.

John made three vocabulary errors in total, namely "make a bad test", "double the year" and "CGSEs."

Marie made three vocabulary errors, "the problem will continue", "those vocab" and "have the teachers better educated".

Dawn made three vocabulary errors, "bad literacy standards", "create your basics", "do a resit about that".

Liz made five lexical mistakes, "lift your level of literacy", "compare it to Dutch", "more beneficial to the education", "if you fail for this part", "how badly one class did in that, er, subject".

Jane made thirteen mistakes, "vocab drops really, really low", "oral communications", "lost the train of thought", "they had black-out", "I don't know how you are when you are in fluency exams", "like a gradual moments", "how big classrooms are", "standardised testings", "the levels of education in Sweden are high", "I was a rather good student", "didn't get higher grades previously", "give that much effort", "was being graded on".

Lena made eight mistakes: "she can do really good", "we can talk really well", "everyone has a start with the same standards", "I can agree with that, but…" , "pass my pronunciation", "he noticed from everyone", "the standard of teaching", "group eight".

Anna made twelve mistakes: "they are updated", "object towards it", "better education of teaching", "qualified in doing their job", "increases the feeling of failure inside the student", "they don't have that level of high concentration", "I also had the idea that", "doesn't have the expected standard after primary school", "below the standard", "high pressure workload", "if that would've been the complete basis of which secondary school you would attend", "how can it really be their fault when it's more inside of the education".

Celia made two lexical errors. "they just go on with the usual years they do it" and "get good teachers in front of classrooms".

Two participants caught themselves making a vocabulary error and quickly corrected themselves. John corrected "compared to" to "compared with", and Dawn corrected "process" and said "progress" instead. I did not count these instances as errors because the students corrected themselves.

Propositional Precision

CEFR Scale Propositional Precision

	PROPOSITIONAL PRECISION
C2	*Can convey finer shades of meaning precisely by using, with reasonable accuracy, a wide range of qualifying devices (e.g. adverbs expressing degree, clauses expressing limitations).* *Can give emphasis, differentiate and eliminate ambiguity.*
C1	*Can qualify opinions and statements precisely in relation to degrees of, for example, certainty/ uncertainty, belief/doubt, likelihood, etc.*
B2	*Can pass on detailed information reliably.*
B1	*Can explain the main points in an idea or problem with reasonable precision.* *Can convey simple, straightforward information of immediate relevance, getting across which point he/she feels is most important.* *Can express the main point he/she wants to make comprehensibly.*
A2	*Can communicate what he/she wants to say in a simple and direct exchange of limited information on familiar and routine matters, but in other situations he/she generally has to compromise the message.*
A1	*No descriptor available*

(Council of Europe, 2001, p. 129).

The scale Propositional Precision is concerned with passing on information in great detail with the help of adverbs expressing degree, describing opinions precisely. Students can be rated from A2 to C2. There is also an additional B1+ level. The A1 descriptor is not available.

Table 8: Degree adverbs used by participants 1

John	Marie	Dawn	Liz
Too (2x)	Very (4x)	Very (2x)	Very (4x)
Very (2x)	Just (8x)	Really (4x)	Really (3x)
Quite (2x)	Really (2x)	Too (3x)	Just (6x)
Just (12x)	Less (1x)	Definitely (1x)	Quite (1x)
Well (1x)	Indeed (1x)		Miserably (1x)
Definitely (1x)	Too (1x)		Well (1x)
Probably (1x)			Definitely (1x)

Table 9: Degree adverbs used by participants 2

Jane	Lena	Anna	Celia
Just (12x)	Too (1x)	Really (15x)	Just (4x)

Really (10x)	Really (23x)	Probably (1x)	Really (1x)
Probably (1x)	Just (26x)	Simply (1x)	Definitely (1x)
Rather (1x)	Quite (1x)	Just (12x)	Very (1x)
Completely (2x)	Well (2x)	Very (2x)	Well (1x)
Ridiculously (1x)		Too (4x)	
Certainly (1x)		Purely (1x)	
Very (1x)		Indeed (1x)	
		Completely (1x)	
		Definitely (1x)	

Table 10: Expressing Opinions: Pair One

John	Marie
"That is my opinion as well" (1x)	"I think…" (13x)
"I personally don't mind it" (1x)	"I'm in favour of" (1x)
"I think" (13x)	"I (also) thought" (2x)
"I do think" (4x)	"I'm not sure if" (1x)
"the way I read it is that" (1x)	"I do think" (1x)

Table 11: Expressing Opinions: Pair Two

Dawn	Liz
"I think…" (23x)	"I think…" (16x)
"But on the other hand" (1x)	"I'm not sure if I agree" (1x)
	"I agree with that" (1x)

Table 12: Expressing Opinions: Pair Three

Jane	Lena
"I'm more inclined to go with" (1x)	"I think…" (23x)
"I think…" (6x)	"I do think…" (1x)
"I do think…" (1x)	"I can agree with that" (1x)

Table 13: Expressing Opinions: Pair Four

Celia	Anna
N.a.	"I think…" (11x)
	"I do think" (1x)

Analysis and Discussion

In this chapter, I will analyse and discuss each scale separately and will consequently answer my research question. I also make some suggestions for further research at the end of the chapter.

Spoken Fluency

Every discussion stayed mostly within the scope of the article except for the discussion between Jane and Lena. They mostly talked about their own experiences and thoughts on university rather than the issue discussed in the article. Moreover, their discussion was also quite informal. This might have helped their fluency as they were talking about familiar topics and issues. Similarly, participants who were not at ease or familiar with the topic discussed in the article might have been less fluent than in other situations. Nevertheless, I judged them on the data that I gathered from these discussions.

Anna and Celia, and John and Marie had comparable speech rates of 152 and 150 words per minute, though there was a striking difference. John and Marie contributed equally to the conversation, which was not the case with Anna and Celia. Celia hardly contributed to the conversation as she only spoke 447 words in 11 minutes and 45 seconds. Anna, however, dominated the conversation and consequently had highest number of spoken words, namely 1162. Anna also had the highest word count of all participants. Only Jane and Lena had similar word counts, with 1059 and 1055 respectively. John, Marie, Dawn and Liz all had fairly similar word counts, though John had the second lowest amount of spoken words: 687. The other three participants had word counts of 843, 803 and 744 words respectively.

Anna's high word count was also reflected in the amount of sentences. She strung 86 sentences together with an average of 13.5 words per sentence. Even though Anna had the highest sentence count, she does not have the highest word average per sentence. Celia had the lowest amount of spoken sentences, 42, with an 10.6 word average. While Jane and Lena had similar word counts, their number of spoken sentences was not. Lena spoke 82 sentences with an average of 12.8 words per sentence, while Jane had 68 with an average of 15.5 words per second, which was the highest average of all participants. John had a comparable amount, 68, though on average, Jane spoke 5.7 words per sentence more than John did, as he had an average of 10.1 words per sentence, which was the lowest average. Marie, Dawn and Liz all spoke 61 sentences, with 13.8, 13.1, and 12.1 word averages per sentence.

While Celia and John had the lowest word counts, they were not equally hesitant. John had a total of 28 hesitations and Celia had zero. The three participants who had the highest word counts, Jane, Lena and Anna, were not equally hesitant as well. Even though

Anna had the highest word count, she had the second lowest number of hesitations. Strikingly, Jane, who had the second highest word count, was the most hesitant of all participants. Lena also had relatively low number of hesitations, namely 8. After Jane, Liz, Dawn, and John were the most hesitant.

Even though Anna was the least hesitant, she had the highest number of restarts, namely 35. Jane had the second highest number with 31. Lena, similarly, had a relatively high number of restarts with 22 restarts in total. After Lena, Marie made the most restarts with 17 in total, followed by Dawn and Liz with 13 and 10 restarts respectively. Celia had the lowest amount of restarts with four in total, followed by John who made nine restarts. Interestingly, both spoke the least of all participants, but also did not need to make a lot of restarts.

There are several ways of assessing a student's fluency. According to Weir (1990), a student should be judged on their communicative skills during a task. He and Young (1998) are in favour of comparing a learner to a native speaker. Hesitation could also be looked at, which Hilton (2014) supports, but Fulcher (2004) is against because he argues that native speakers can be naturally hesitant as well.

For this research, hesitation and speech rate are chosen as fluency markers as I am interested in their levels compared to native speakers, but rather the differences among the students themselves. Despite the fact that hesitation as a proficiency marker is heavily disputed, hesitation can still be a useful marker in determining who is more fluent when it is analysed in combination with speech rate.

There is evidence that the three fastest speakers have the most restarts of all participants. As the duration of the hesitations is influenced by speech rate, both factors need to be taken into consideration when judging who is more fluent than the other. Jane, for instance, has the second highest word count of all participants and is thus one of the fastest speakers, but also has the highest number of hesitations and the second highest number of restarts. On paper, she would be the least fluent of all participants looking solely at restarts and hesitations, but because she is a very fast speaker, hesitations and restarts are hardly noticeable during the discussion. A similar analysis can be drawn for Anna and Lena, who, too, are a very fast speakers, make a lot of restarts, though do have a lower number of hesitations. Despite the fact that they are hesitant in terms of numbers, their hesitations do not hinder their speech and communication. They are also able to hold the floor for longer stretches of time despite their hesitations. They are thus fluent because they are able to keep their conversation going and simultaneously hold the floor for a long time without drawing attention to their hesitations since they were so short.

EL, Anna and Lena are therefore rated with C2/C1 CEFR rating as they are not overtly C1 or C2, but in between. C2 is not the most fitting as they are hesitant and make a lot of restarts. Their speech does come across as natural and effortless and thus fit the C1 descriptor perfectly.

Liz, Marie, and Dawn are rated with a C1. They are not particularly slow or fast speakers, or overly hesitant. This means that their hesitations were more noticeable than Jane, Lena, and Anna's, though they still produce more words than the two slowest speakers. Their speech is also spontaneous, and they are able to articulate their thoughts quite fast without long pauses. They are also able talk for a longer stretches of time.

John is rated with a B2+ level. John, who is a relatively slow speaker, is the third most hesitant speaker, though does not make a lot of restarts. Because of the fact that John is a slow speaker, hesitation is more apparent in comparison with Jane, Lena and Anna. John thus comes across as a less fluent speaker.

Celia is rated with a B2. Assessing Celia is difficult, though, as her word count is the lowest of all participants simultaneously is the least hesitant and makes the least restarts. In her case, hesitation does not play a role, but because she does not produce a lot of words, there is not a lot of data to judge her on. It seems that Celia thoroughly thinks about what she wants to say before speaking and is not able to hold the floor for very long. Consequently, she might be less inclined to hesitate or to make a restart because she has thoroughly thought about what she is going to say before doing so.

In conclusion, assessing fluency is difficult and multiple factors need to be taken into consideration. For this research, hesitation, the number of restarts and number of produced words are taken into consideration. The students who produce the most words, turn out to be the most hesitant and/or make the most restarts. Though because of their speech rate, their hesitations and restarts are not as noticeable as students who speak slower. The duration of the hesitations made by the fast speakers is shorter than those who speak slower. The fast speakers are able to go on more quickly than the slower speakers, which makes their speech more effortless than the slower speakers. There is also a group between the fast speaker and the slower speakers, who was neither overly hesitant nor made a lot of restarts. One student does not produce a lot of words, is not hesitant or makes a lot of restarts. Though, she is not able to hold the floor for longer stretches of time, which makes her also less fluent. Jane, Lena, and Anna are rated C2/C1; Liz, Dawn, Marie are rated with a C1, John is rated with B2+ and Celia with a B2.

Vocabulary Range

For this scale, the participants' lexical density, type-token ratio and the number of matches with the vocabulary of the article were researched.

John's vocabulary had the highest lexical density with 48.03%, followed by Celia with 44.51%. Dawn had the third highest percentage with 43.96%. Liz and Jane has comparable percentages, 43.54% and 43.05%. Marie's percentage was 42.34%. Lena and Anna had the lowest lexical density percentage, with 40.11% and 39.58%.

Celia had the highest type-token ratio with 0.39, followed by John with 0.32. Dawn had the third highest percentage with 0.30. Liz, Marie, Jane and Lena had a type-token ratio 0.27. Anna had the lowest type-token ratio: 0.26.

Dawn's vocabulary matched most with the article with 115 matches in total and a percentage of 47.52%. Even though Lena and Anna had the most matches after Dawn with 104 and 112 matches, their percentages are not the highest as they are 36.24% and 36.96%. Liz has the second highest percentage with 46.23% and 92 matches. Despite having the lowest number of matches, namely 77, Celia had the third highest percentage with 43.26%. Marie and John had 96 and 91 matches and percentages of 42.68% and 40.81%. Jane had the lowest percentage of all participants, 34.26%, despite the relatively average number of matches, 99.

John, Celia, and Dawn have the highest type-token ratios and lexical density percentages. Dawn also has the most article vocabulary matches in total and the highest percentage as well. Dawn thus has a wide vocabulary range, especially in her vocabulary use similar to the article, but her number of unique words was also fairly high. Her type-token ratio and lexical density also support this. John and Celia have high percentages and type-token ratios, though have a lower score in their number of vocabulary matches as they have the two lowest numbers of matches. Looking at Celia's number of unique words, it is evident that she has not demonstrated a very wide range of vocabulary during the discussion. John has the second lowest number of matches, but his overall number of unique words is not the lowest of the participants.

Anna has the second highest number of vocabulary matches but also the most unique words with 303 in total. Anna, however, does have the lowest type-token ratio and lexical density. Evidently, the ratio and lexical density do not necessarily line up with the number of unique words. In spite of Anna's relatively low type-token ratio and lexical density percentage, she has demonstrated a wide range of vocabulary, both article-related and in terms

of unique words. A similar conclusion can be drawn with Lena's data. She, too, has a lot of vocabulary matches, though a relatively low type-token ratio and lexical density.

Jane, Marie and Liz have comparable vocabulary matches, namely 99, 96, and 92. Their type token ratios are identical: 0.27. Their lexical density percentages are also fairly close together. Their number of unique words are, however, not comparable. Jane has the most with 289, which is the second highest of all participants. Liz has 199 and Marie 225. Despite their fairly comparable data, the number of unique words showed that Jane has a wider range than Liz and Marie.

 I rate Dawn, Anna, Jane, and Lena with a C1. Their data is fairly similar as they all had a large number of unique words and have the most article vocabulary matches. Liz, Marie, and John are rated with a C1/B2. They have fewer article vocabulary matches than the aforementioned participants, though not significantly fewer to rate them with a B2. I rate Celia with a B2, despite her high type-token ratio and lexical density percentage, because her range is relatively narrow and her number of vocabulary matches is the lowest.

In conclusion, some students demonstrate a wider vocabulary range than other students judging from their unique words count and number of article vocabulary matches. Type-token ratio and lexical density often do not line up with the number of unique words and vocabulary matches, except for one case, namely Dawn. Along with Anna, Jane and Lena, Dawn was rated with a C1. Liz, Marie and John were given a C1/B2 and Celia a B2.

Vocabulary Control

All participants made lexical errors, though some more than others. The lexical errors could be divided up into collocation errors, article-related errors, pronoun-errors, and single/plural-related errors.

John mostly made collocation errors. "[M]ake a bad test" should be 'do badly on a test', and "double the year" should be 'retake the year'. John also mistook 'GCSEs' for "CGSEs".

Marie used the wrong pronoun with "those vocab", which should be 'that' since vocab is a single noun. She also made a collocation error, namely "the problem will continue", which should be something along the lines of "the problem will still not be solved". "[H]ave the teachers better educated" is also incorrect English. A better alternative would be 'improve the education of teachers'.

Dawn made three collocation errors. "[B]ad literacy standards" should be 'bad literacy levels'. "[C]reate your basics" should be 'master your basics'. For "do a resit about that" should be without 'about that'.

Liz mostly made collocation errors. "lift your level of literacy" should be 'improve your literacy level' and "compare it to Dutch" should be 'compare it to the Dutch system' or 'compare it to the Netherlands'. "[M]ore beneficial to the education" should have been followed with a preposition and a noun because the definite article is out of place in this phrase. "[I]f you fail for this part" should be 'if you fail this part'. Finally, the preposition "in" in "how badly one class did in that, er, subject" should be 'at'.

Jane made multiple errors. Three errors were single/plural-related errors. "[O]ral communications" should be 'communication'. "Like a gradual moments" is incorrect since the article does not collocate with the plural noun and is therefore ungrammatical. "[S]tandardised testings" is also incorrect because 'testing' as a nominalisation can only be used in single form. Jane also made several collocation errors. 'Make such an effort' is the right expression for "give that much effort" . "[W]as being graded on" is incorrect. 'Get a grade for' is a more suitable expression. "[D]idn't get higher grades previously" is incorrect because "previously" should be 'before'. "[T]he levels of education in Sweden are high" is also incorrect. A better alternative would be 'the quality' instead of 'level'. Jane also made three article-related errors. "they had black-out" should be with an indefinite article or be verbalised: 'they blacked-out' . "[L]ost the train of thought" should be with a personal pronoun instead of a definite article. In "I was a rather good student" the indefinite article should be before the adjective. Lastly, Jane made some other lexical errors. "[V]ocab drops really, really low" is wrong use of vocabulary. She intended to say that her use of vocabulary during exams is not really up to standard. "how big classrooms are" is wrong in this context as it should be "how big classes are" since she meant the group of people and not the actual room. "I don't know how you are when you are in fluency exams" is an awkward sentence despite its grammaticality. She meant something along the lines of 'how you behave or act during a fluency exam'.

Lena made several collocation errors. "[H]as a start" is incorrect and can be replaced with 'starts'. "[P]ass my pronunciation" is an incomplete phrase because 'exam' should come after "pronunciation". "[T]he standard of teaching" should be 'teaching standard'. Lena also made some incorrect phrases. "We can talk really well" and "I can agree with that, but…" are grammatical, though, awkward sentences because of the use of 'can'. In "She can do really good", an adjective was used instead of an adverb, which renders the sentence ungrammatical

and incorrect. The over-use of 'can' could be caused by L1 transfer because of the Dutch equivalent 'kan'. She also referred to 'year eight' as group eight, which is also caused by L1 transfer. "[H]e noticed from everyone" is also incorrect English, but Lena meant to say that the teacher had observed the children correctly in year one.

Anna made several kinds of errors. "They are updated" is incorrect since 'they' refers to teachers and they cannot be updated. "[T]hey don't have that level of high concentration" is also incorrect word use in this context. 'They don't have a lot of concentration' would be a better alternative. "[I]ncreases the feeling of failure inside the student" is incorrect and should be something along the lines of 'increases the student's sense of failure'. Anna also made four collocation errors. "[O]bject towards it" should be 'object to', "qualified in doing their job" should be 'qualified for doing their job'. "[D]oesn't have the expected standard" should be 'doesn't meet the expected standard'. She made two article-related errors. "[B]elow the standard" should be without the definite article. In "how can it really be their fault when it's more inside of the education" "education" should be without the definite article. Moreover, the sentence is worded quite awkwardly and could be improved by saying 'when the fault lies within the educational system'. "[H]igh pressure workload" is also incorrectly phrased and should be 'high pressure' or 'workload'. Anna also made a lexical error which is still grammatical but not entirely correct in English, namely "I also had the idea that". This is an error caused by L1 transfer from Dutch.

Celia made one collocation error, "get good teachers in front of classrooms", which should be 'classes'. Celia also made one awkward sentence, "they just go on with the usual years they do it". She meant to say that children go on to secondary school even after failing their resits, but was not able to formulate that properly.

Thus participants mostly made several kinds of collocation errors. These errors did not hinder their communication. In all cases, the conversation went on without the interlocutor having to ask for clarification. Their collocation errors were mostly related to incorrect use of prepositions. Strikingly, two participants made an identical lexical error. Jane and Celia both made the mistake of confusing 'classrooms' for 'classes'.

Several students strung some sentences together which were not necessarily ungrammatical, though still incorrect. Anna's "how can it really be their fault when it's more inside of the education" is an example of awkward, though still grammatical, phrasing. Despite the fact that the sentence is incorrect English, Anna still got her point across.

Two students made errors that could be caused by L1 transfer from Dutch. Anna said "I also had the idea that" which is a grammatical sentence though not quite a typical English

sentence. Lena used 'can' several times, like in "I can agree with that, but…". Again, it is not an ungrammatical sentence, though very atypical as 'I agree' would be more straightforward and correct. Lena also said 'group eight', which is a literal translation from 'groep acht' in Dutch, instead of 'year eight'. This is also an error caused by L1 transfer.

There seems to be a link between speech rate and the number of errors made. Anna, Jane, and Lena have the highest word counts of all participants and also have made the most errors with twelve, thirteen and eight errors respectively. Similarly, Celia has the lowest word count and also has the lowest number of errors made. Though it needs to be said that even though John has the second lowest word count, he has made three mistakes just like Marie and Dawn, both of whom have higher word counts.

Liz, Anna, Lena and Jane make the most lexical errors of all participants. Therefore I thought that C1 does not apply to them as some significant errors were made by these participants. B2, then, would be more fitting as their word use was mostly correct, though significant mistakes can still be spotted.

For John, Celia, Dawn, Marie, C1 is most fitting. They make the least mistakes which are often collocation errors. C2 is not fitting because their English was not error-free.

In conclusion, there is a clear divide between participants that made fewer errors and more errors. The students with more errors are consequently rated lower on the CEFR scale than the other participants. Generally, the participants' errors are mostly collocation errors that do not hinder communication at all. This can be explained because the all participants were Dutch and not native speakers. If the interlocutors had been native speakers, the errors might have caused hindrance during discussion. Some participants created some awkward phrases that were not perfect English, though like the collocation errors, these sentences have not hindered their discussion and were able to go on without any difficulty.

Propositional Precision

Adverbs

Anna used the most degree expressing adverbs with 10 in total. Jane followed with 8 in total. John and Liz had 7, Marie had 5. Dawn, Celia and Lena had the least with all 4 in total.

All participants except Dawn used 'just' multiple times. 'Really' was used by every participants except John. 'Very' was also a popular adverb as it was used by every participant except for Lena. Five out of eight participants used 'too' as an adverb. Five participants, Celia, John, Anna, Dawn, and Liz used 'definitely'. Half of the participants used 'well'.

'Probably' was used by Anna, EL and John. Anna and Jane both used 'completely'. Marie and Anna used 'indeed'.

Jane was the only one to use 'rather', 'ridiculously', and 'certainly'. Anna uniquely used 'simply' and 'purely'. Marie was the only one to use 'less'. Liz was the only participant to use 'miserably'.

Generally, there is not a lot of variety in the use of degree adverbs. Jane and Anna use of degree adverbs is the most varied compared to the other participants. Marie and AB's also have demonstrated some unique adverbs, while the rest use of degree adverbs is made up common adverbs that other participants also used. Adverbs that are used by nearly every participant, like 'just' or 'really', are used multiple times by the individual participants as well. The use of degree adverbs is mostly repetitive and not varied.

Expressing opinions

Everyone except Celia used 'I think' when they expressed their opinion. John, Marie, Jane and Lena used 'I do think' as well. Marie also used the past form, "I thought". Liz and Lena both expressed their opinions with 'agree'. Marie and Liz used the 'I'm not sure if…" construction.

Celia expressed her opinion without any particular expressions. John was the only one to use these three phrases, "[t]hat is my opinion as well", "I personally don't mind it", "the way I read it is that". Marie was the only participant to use "I'm in favour of". JH uniquely used "but on the other hand". Jane was the only on to use "I'm more inclined to go with".

Participants thus mostly use the same construction, 'I think', when expressing their opinion. They also use this phrase often during their discussions as can be seen from the results. 'I think' clearly is an expression that is the easiest for participants to use and they consequently tend to over-use the phrase. Some participants, however, use other constructions.

John used three other phrases besides 'I think', and was one of two participants to express their opinion in such a diverse way. Marie also used three other phrases. The other participants way s of describing their opinions were rather limited. Interestingly, while Celia did express her opinion during the discussion, she did not use any specific constructions to formulate her opinion. She stated her opinion as if it were a fact without any tentativeness.

As Jane and John are the most varied in their use of degree adverbs and expressions, I rate them with a C1. I rate Anna and Marie with C1/B2, as they were neither C1 or B2, but a bit of both as they did use various adverbs and expressions. I thought B2 is more fitting for Liz, Dawn, Marie and Celia. Their use of degree adverbs and expressions is less diverse than

the aforementioned participants. Even though Celia does not use specific constructions in expressing her opinion, she still expresses her opinion. That is why I rate her with a B2 level as she was still able to get her point across.

In conclusion, some degree adverbs and the phrase 'I think' are used, perhaps even over-used, by all participants. A few participants are able to express their opinions with various degree adverbs and different constructions, while others ability to do so was rather limited.

Answering Research Question

My research question was: how different are the oral proficiency levels of second-year students of English at Radboud University? My hypothesis was that would be a clear distinction between advanced and less advanced students. The students who were more fluent were expected to make fewer errors and restarts, to demonstrate a more varied use of vocabulary, and were also expected to be less hesitant than the students who were less fluent.

My hypothesis was not in line with the results. Based on my data collected from the discussions, a general clear-cut divide between the participants cannot be easily made. The participants can be divided per scale, but not in general. The results show that it is not true that if one participant is more fluent than others, he or she is also less hesitant, makes fewer restarts or mistakes. There is evidence that the opposite is actually true. Participants that were less fluent made fewer mistakes and restarts, and were also less hesitant than the more fluent participants. Fluency and vocabulary range were also not completely mutually exclusive as one student who I judged as moderately fluent, had one of the widest vocabulary ranges. Similarly, more fluent students did not necessarily use more degree adverbs or used varied phrases when formulating their opinions.

Despite the fact that my hypothesis was incorrect, some conclusions can be drawn from the data discussed. There are some differences between these eight second-year students' oral proficiency levels. Some students are more fluent than other students. I rated the students ranging from B2 to C2/C1. Every participant made lexical mistakes, which were often collocation errors. Some participants, however, made more errors than others and these errors were usually not collocation errors. Even though the participants made some mistakes, it did not hinder their communication at all. The participants were rated with C1 and B2. Some students used more vocabulary from the article they discussed. Similarly, students also differed in their number of unique words used. Students thus had different vocabulary ranges and were rated with C1 and B2. The participants were similar in their use of degree adverbs

and the way they expressed their opinion. Some students used a few other expressions or adverbs, though still frequently used certain adverbs that were also used by the other participants. They were rated from B2 to C1.

Thus, the second-year students I researched do have different proficiency levels. I rated the students from B2 to C2/C1, so there certainly were some differences in proficiency levels. However, not a single student was rated with a C1-level, for instance, on every scale. This means that the oral proficiency levels among students is different, but also the ratings of the students themselves are not necessarily identical.

Table 14

CEFR Assessments

Participants	Spoken Fluency	Vocabulary Range	Vocabulary Control	Propositional Precision
John	B2+	C1/B2	C1	C1
Marie	C1	C1/B2	C1	C1/B2
Dawn	C1	C1	C1	B2
Liz	C1	C1/B2	B2	B2
Jane	C2/C1	C1	B2	C1
Lena	C2/C1	C1	B2	B2
Anna	C2/C1	C1	B2	C1/B2
Celia	B2	B2	C1	B2

Suggestions for Further Research

Conclusions drawn from this research apply to the discussed participants only. Like Davis (2009) argues, my findings are not all-conclusive and apply only to the discussions I recorded at the time. Doing a different task, like an interview, in another setting might yield different results. My research thus can only give some indications to how different the oral proficiency levels of second-year students of English are. It can, however, serve as a starting point for further investigation into the oral proficiency levels of second-year students. The data I collected can be investigated even more elaborately. Spoken fluency and hesitation could be looked into by measuring how much time is taken up by students' hesitation. Vocabulary range could also be further investigated by looking at expressions used by students. It is also a possibility to research more second-year students to gain more insight in this group's oral proficiency. Studying more students will naturally give more support and weight to possible conclusions or evidence.

The students could also be compared to native speakers. I used hesitation and speech rate to assess the students' spoken fluency, but it is also a possibility to investigate how similar or dissimilar their speech is to that of a native speaker's.

This research may also be expanded by using other CEFR scales as intensively in order to get a more elaborate view of second-year students' proficiency. Like North (2014a) states, the CEFR is a good starting point for assessment which has turned out to be an effective way to use the Framework for assessment. It is important for future users of the Framework to actively engage with the scales and its descriptors and adapt it to their needs. This way the Framework can be used to its full potential.

Another possibility could be to investigate first- or third-year students, or even master students, and compare them to the second-year students. This might yield interesting results and conclusions as to how students progress during their studies at Radboud University.

Conclusion

This research dealt with the following research question: How different are the oral proficiency levels of second-year students of English at Radboud University? My hypothesis was that would be a clear distinction between advanced and less advanced students. The students who were more fluent were expected to make fewer errors and restarts, to demonstrate a more varied use of vocabulary, and were also expected to be less hesitant than the students who were less fluent.

In order to answer the research question, eight second-year students' of English oral proficiency levels were studied. They each had a ten-minute discussion with a fellow student of a comparable level. Every discussion was transcribed orthographically and then further analysed with CorpusTools and Antconc. The CEFR scales Spoken Fluency, Vocabulary Range, Vocabulary Precision and Propositional Precision served as rating scales, but also as starting point for in-depth analysis of the students' oral proficiency levels.

The results showed that my hypothesis was incorrect. Students who were more fluent did not necessarily make fewer mistakes and restarts, nor were they less hesitant. Evidence was found that it was the opposite. Students who were less fluent made fewer lexical mistakes, made fewer restarts, and were not as hesitant. Though it must be said that speech rate influenced how noticeable hesitation was. More fluent students spoke faster which made hesitation less noticeable than those who spoke slower. Some students also had a wider vocabulary range than the others, though a wider range was not linked to being more fluent. Next to these differences, there were also similarities. All participants made lexical mistakes that were often collocation errors. Even though the students made errors, their communication was not hindered by their mistakes. The students also used a limited number of degree adverbs, and often over-used certain adverbs. They also had similar ways of expressing their opinions.

Students are rated from B2 to C2/C1, but not a single student is rated with one particular level on all scales. This means that students' proficiency cannot be easily be labelled with one level and that their skills should be researched and rated separately. The results have shown that there were differences in the students' oral proficiency levels, but also that being more fluent does not mean that other skills are just as advanced. These findings have given some insight into the oral proficiency levels of second-year students, but there are still many ways that could complement this research in order to get a more complete view of how different the oral proficiency levels of second-year students are.

References

Adams, Richard. 7 April 2015. Tories to announce resits for pupils who fail end of primary school exams. The Guardian. Retrieved from: http://www.theguardian.com/

Anthony, L. (2014). AntConc (Version 3.4.3) [Software]. Tokyo, Japan: Waseda University. Available from http://www.laurenceanthony.net/

Alderson, J. C. (2007). The CEFR and the need for more research. *The Modern Language Journal*, 91(4), 659-663.

Alderson, J. C., Figueras, N., Kuijper, H. Nold, G. Takala, S. & Tardieu, C. (2006). Analysing tests of reading and listening in relation to the Common European Framework of Reference: The experience of the Dutch CEFR construct project, *Language Assessment Quarterly*, (3)1, 3-30.

Council of Europe. (2001). *Common European Framework of Reference of Languages: Learning, Teaching, Assessment*. Cambridge: Cambridge University Press.

Davidson, F. & Fulcher, G. (2007). The Common European Framework of Reference (CEFR) and the design of language tests: A matter of effect. *Language Teaching*, 40(3), 231 – 241.

Davis, L. (2009). The influence of interlocutor proficiency in a paired oral assessment. *Language Testing*, 26 (3), 367–396.

Edmonds, A. & Leclercq, P. (2014). How to assess L2 proficiency? An overview of proficiency assessment research. In P. Leclercq, A. Edmonds, and H. Hilton (eds.), *Measuring L2 Proficiency: Perspectives from SLA* (pp. 3-23). Bristol: Multilingual Matters.

Esteki, B. (2014). The relationship between implicit and explicit knowledge and language acquisition. *Theory and Practice in Language Studies*, 4(7), 1520-1525.

Fulcher, G. (2004). Deluded by Artifices? The Common European Framework and harmonization. *Language Assessment Quarterly*, 1(4), 253-266.

He, A.W. & Young, R. (1998). Introduction: assessing second language speaking ability. In R.Young and A.W. He (eds.), *Talking and Testing* (pp. 1-26) Amsterdam: John Benjamins.

Hilton, H. (2014) Oral fluency and spoken proficiency: considerations for research and testing. In Pascale Leclercq, Amanda Edmonds, and Heather Hilton (eds.), *Measuring L2 Proficiency: Perspectives from SLA* (pp. 27-53). Bristol: Multilingual Matters.

Hulstijn, J. (2007). The Shaky Ground Beneath the CEFR: quantitative and qualitative dimensions of language proficiency. *The Modern Language Journal*, 91(4), 663-667.

Little, D. (2006). The Common European Framework of Reference for Languages: Content, purpose, origin, reception and impact. *Language Teaching*, 39(3), 167-190.

Little, D. (2007). The Common European Framework of Reference for Languages: perspectives on the making of supranational language education policy. *The Modern Language Journal,* 91(4), 645-655.

Little, D. (2011). The Common European Framework of Reference for Languages: A research agenda. *Language Teaching,* 44(3), 381 – 393.

North, B. (2000). *The Development of a Common Framework Scale of Language Proficiency*. New York: Peter Lang.

North, B. (2007). The CEFR illustrative descriptor scales. *The Modern Language Journal,* 91(4), 656-659.

North, B. (2014a). *The CEFR in practice*. Cambridge: Cambridge University Press.

North, B. (2014b). Putting the common European framework of reference to good use. *Language Teaching*, 47(2), 228 – 249.

Ross, S. (1992). Accommodative questions in oral proficiency interviews. *Language Testing*, 9(2), 173 – 185.

O'Donnell, Mick. (2007). UAM Corpus Tools (Version 2.8.14) [Software]. Available from: http://www.wagsoft.com/CorpusTool/

Osborne, J. (2014) Multiple assessments of oral proficiency: evidence from a collaborative platform. In P. Leclercq, A. Edmonds, and H. Hilton (eds.), *Measuring L2 Proficiency: Perspectives from SLA* (pp. 54-70). Bristol: Multilingual Matters.

Saville-Troike, M. (2012). *Introducing second language acquisition*. Cambridge: Cambridge University.

Weir, C. J. (1990). *Communicative language testing*. Hertfordshire: Prentice Hall International.

Zhang, Runhan. (2015) Examining the role of implicit and explicit L2 knowledge in general. *International Journal of English Linguistics*, 5(3), 79 – 87.

Appendix A: Guardian Article
Tories to announce resits for pupils who fail end of primary school exams
By Richard Adams

Policy would apply to 100,000 11-year-olds who begin secondary school with inadequate levels of literacy and maths skills

Children who fail exams at the end of primary school will have to resit them the following year, in an effort to ensure pupils are "secondary ready", the Conservatives are to announce as part their general election campaign push on education.

The resits would apply to about 100,000 11- or 12-year-olds in England who leave primary school with inadequate levels of literacy and maths skills, based on their performance in key stage two standardised exams.

Under the new Tory policy, from September 2016 pupils deemed to have failed to reach the required levels will have two chances to sit retakes, administered by their secondary schools. The policy will be outlined by the education secretary, Nicky Morgan.

But critics said the new tests would be redundant and could act as a powerful disincentive against state secondary schools taking pupils with weak key stage 2 results because of the extra workload and pressure on results.

But Morgan said the aim of the new test was to stop children being "written off" too early by poor results at primary school.

"We know that the biggest predictor of success at GCSE is whether young people have mastered the basics at age 11. That means if we fail to get it right for young people at the start of secondary school they'll struggle for the rest of their time in education," she said.

"Under Labour, one in three children left primary school unable to read, write and add up properly. Thanks to our reforms and teachers' hard work, we've seen that fall to just one in five.

"But even one child falling behind, or being written off, is a child too many. That's why the next Conservative government will require schools to enter any child who doesn't have the literacy and numeracy skills they need to succeed in secondary school for new year seven resit tests that will guarantee they've caught up."

Tristram Hunt, the shadow education secretary, said the policy was "a desperate attempt" to distract from Conservative failures on school standards.

"On their watch, 1.6 million pupils are being educated in schools that are less than good. And, as a result of David Cameron's unqualified teachers policy, more than 400,000 pupils are being taught by unqualified teachers," Hunt said.

"Labour has a better plan for education. We will ensure that every teacher is qualified or working towards qualified teacher status, and introduce a new master teacher status to raise the standing of the teaching profession. That is how we improve education for every child in every classroom."

Christine Blower, the general secretary of the National Union of Teachers, said the resits would be totally unnecessary and at odds with government policy giving autonomy to schools.

"It is simply hard to see – other than an obsession with testing and metrics – what possible value there is in doing this," she said.

Russell Hobby, the general secretary of the National Association of Head Teachers, said the new policy was "a threadbare approach, that 'if it moves, test it'," adding: "This policy has confused teaching and testing, and thinks that tests improve teaching."

A Conservative spokesman said it was expected that 80% of those sitting the resits in English and maths would pass, and that those who continued to fail would not be held back to repeat year seven.

Failure rates in the resits would be a "red flag", he said, with the data published as part of school performance tables. Ofsted inspectors will be expected to use the results as part of school assessments.

"There is no job that doesn't require English and maths and this is about making sure every child gets the best start in life and that our country can compete in the world," David Cameron said, backing the new tests.

The resits set by the Department for Education's standards and testing agency would be a "slimmed down" exam of year six English and maths but would include some material from the year seven secondary school courses. Pupils taking resits would still be taught in their usual year group.

Schools will be able to enter students twice if needed for the new tests, in either spring or summer term of year seven. Pupils with special educational needs that prevented them from achieving the required standard will be exempt.

Secondary schools currently receive £500 in catchup funding per pupils for those who arrive without having met the expected standard in year six.

"We recognise that secondary schools have to do a lot of work to pick up the results of some of the less able children coming from primary schools," the Conservative spokesman said.

Of the 100,000 children who fail to reach the expected standard in English and maths at 11, only 7% went on to gain five good GCSEs including those two subjects, compared with 72% who made the grade at the end of primary school, according to DfE figures.

Appendix B: Transcription 1
/   Said simultaneously
-                        Interruption
--   Restart
+   Partial word
(( ))   Unintelligible speech

<u>Pair One</u>

Marie   So, what do you think of this--this announcement?

John   Erm, well, actually, I think it is not too bad of an idea. Erm. I mean it does have its downsides, like, getting stuck for another year, erm, at primary school, but I don't think that it's -- that the, erm, the disadvantages outweigh the advantages, to be honest.

Marie   Well, what advantage would it have to, like -- if they fail the test, they have not had the level that they should have after six years. How would one test make sure that they do get that level with one extra year?

John   Yeah, that's a strange thing, because you could also just make a bad test, you know, not even be bad at, like, maths or English. Erm, maybe you are just stressed or not feeling well or whatever, and you just fail because of that. Erm. And that would be a real shame if they actually have to double the year for that.

Marie   Yeah.

John   Erm.

Marie   I don't know exactly how it works. Do they normally get a resit after they fail the first test or is this, like, the first time they come up with that?

John   Erm. The way I read it is that this is the first time actually that they come up with it.

ME   I think that's the case, yeah.

John   Yeah. If-If you fail it, you just fail it and go on to secondary school.

Marie   That is a bit odd, because then -- why would you do the test? If it doesn't -- If it means that you can just go on?

John   Yeah probably just to see (( )). But, it is not very effective, no. Erm. Well like I said, I do think it is a good idea, but, erm, I'm not sure if they should actually wait a -- wait a whole year to do the resit. I mean, they just -- they just, erm. Do it a couple weeks after, I don't know.

Marie   I think it said somewhere that they could do it both in autumn, I think, er, summer or in spring. Erm. So that would mean, that you can -- when you think you are ready you can do resit.

John   Oh, okay. Still it takes up another year.

Marie   Yeah.

John   For you to go to secondary school.

Marie   I also thought it was very strange that at some point they said that, erm, that the students who fail the test will still go on to secondary school and will still be in the same class as their fellow students, but take extra courses or takes, er, some courses from secondary school and some from primary school so that they, kind of, keep lagging behind.

John   Yeah. So they can catch up a bit, but I don't think that actually works like that.

Marie   No, I mean. Then leave them just one extra year in primary school instead of sending them off to secondary school and.

John   Because of –- yeah, the work will just pile up eventually /and.

Marie   /Yeah, exactly.

John   What is it saying in? Like, "only seven percent went on to gain five good CGSEs", erm. Seven percent compared to -- compared with seventy-two percent. That's quite a lot.

Marie   Yeah, I do see that they have to change something, but I -- I'm not sure if this is the right solution for the problem.

John   Yeah, erm. I don't know. Do you have any ideas how they could actually do this efficiently?

Marie   Well, they said in the article that it was mainly the problem that the teachers were not, er, educated enough to actually teach the children. Ah, yeah -- you could do something with that, I mean that seems like a fair solution to the problem, like, erm, have the teachers better educated instead of, er, the diff- +different, individual pupils, because, er, if the teacher is better then, erm, they can continue to teach pupils better. And if you just focus on the individual pupil then the problem will continue because the teachers are still inadequate.

John   That's very true, yeah. I also think maybe they should do an earlier test to see where the children actually stand, erm, before this one /so if someone has problems --.

Marie   /I do think they do that.

John   Yeah?

Marie   Like, in the Netherlands we also have that. We have this one big test at the end. We also have just, er. The-the, erm, test that can prepare for that, that does not really count, but it is the same, er, format that we can -- can come up -- do what we know it's...

John   To just get used to it.

Marie   Yeah.

John   Well, erm, I don't know how they do that in England, but I think that if, erm, like, a test -- a pre-test let's call it that, erm, shows that a student, ah, has problems, let's say English, they should be, erm, guided a bit more than -- so their level goes up.

Marie   Then that costs extra /money and they can't afford it.

John   /That's true.

John   Yeah, well that's why: "secondary schools currently receive five hundred pounds in catch-up funding." Erm, well, I think you could well spend their money in primary school than...

Marie   Yeah, like, instead of -- giving money to the resits, you can also give the money to.. yeah.

John   I think that would be more efficient /to be honest.

Marie   /Yeah. [laughs] Yeah.

John   'Cause in secondary school you already need more money, because, well like I said, the work is already piled up more, so they have to do more to catch up, actually. So it might be bit harder then... To get them on the right level.

Marie   Yeah and if-if you help them in primary school then they can start fresh at secondary school and, erm, then they can just continue like the rest of the students instead of (( )) having their work piled up because they all, er, are already behind when they enter secondary school.

John   In general, I do think this is a good idea.

Marie   The resits?

John   Yeah. It kind of depends on how they work it out specifically, but, erm.

Marie   Yeah. It is indeed -- the opportunity of a resit, that I think would help. But what they do with the result of the resits and erm, how they -- yeah, I don't think -- in the article it says that they then go on to secondary and take on some subjects of primary school and some of secondary school, I don't think that would work. 'Cause then you don't really -- yeah, you keep lagging behind. But if-if they, er, were to spend an extra year of focussing on the basic skills, like English and maths for example, then I think it would work, because then you would give them an extra opportunity to, er, develop those subjects.

John   So you just mainly focus on those specific ones?

Marie   Yeah, just one extra year of primary school instead of going on, off to secondary school.

John   To-to make sure, well, you're up to -- up to the test.

Marie   Yeah, so you can start fresh at secondary school.

ME   Well, you've talked about these tests, but what do you think of testing in general? Do we do enough testing, or should there be more testing, less testing? You know, if you were to look at university for example. Maybe you can give that some thought.

John   Erm. I think it kind of depends, erm, on the course because for some courses we only have one exam and, well, I personally don't mind it, because you have to study less, but.

Marie Could you give an example of one of those courses? What you're thinking of?

John   Erm. What is it called, again? Ah, I can't remember the name of that.

Marie The erm, like, erm, the phonology course we had? Did we have that? No, we had two of them.

John   No, there were two. Yeah, some of the minor courses, /erm-

Marie /and the core curriculum.

John   And the core curriculum as well, yeah.

Marie And I thought, especially with the core curriculum, it was kind of hard, because you had, like, fourteen weeks of information, but at the same time the tests were easier.

John   Erm, yeah. Depends on the course, I think, but, yeah, okay.

Marie Okay, fair enough.

John   The fail rate was quite high in my course.

Marie Oh.

John   Erm. Yeah, so for those kind of courses I do think it's a bit too, yeah. There aren't enough, ah.

Marie Yeah.

John   Opportunities to test your skills. (( )) I do think once a period is more than enough.

Marie I think in general it's very good, because yeah. I mean. Most of tests come from OCS, but I think that those are things that actually should be tested. Vocab and -- well, vocab a bit less, maybe. But I understand why they want us to learn vocab. I don't know if it's the right way to learn those vocab, but.

John   Well, in OCS they just take a couple of sub- +subjects, well, (( )) together, erm, to test, like, erm. Pronunciation and vocab and...

Marie Yeah, I think those are all necessary to /(( )).

John   /Yeah, they are all just different topics that need to be tested.

Marie I think with Britlit it's very good that we have two, because -- you just have two separate periods, so it's very logical to have just a test after each period.

John   It makes it a bit clearer.

Marie Yeah. Otherwise it would be way too much.

John   Definitely.

ME   So, in the end, you're in favour of the resits that the Tories want to install? Or are you more /against?

Marie /I'm in favour of the resits but I'm against the way they want to, er follow up from the resits.

John   That is my opinion as well.

Appendix C: Transcription 2
Legend:

/	Said simultaneously
-	Interruption
--	Restart
+	Partial word
(( ))	Unintelligible speech

Pair Two

Liz   Erm. So, the question is, is taking those resits erm, more beneficial to the education than better teachers or something?

Dawn Yeah, I think it is. I-I think the, er, critics say that, er, these resits would, erm, be too much testing and they want them to -- or they want the schools to, er, improve the teaching standards and not the erm -- I think, yeah.

Liz   Well, it-it is said in the text that, er, that some teachers are unqualified. Erm. I think that doesn't help, erm, the quality of education, erm, but I do think that someone should be allowed to take a resit if that might improve their chances in secondary school.

Dawn Yeah, I think it's -- I-I agree because, er, it says that, er, when children enter secondary school they are quickly written off, er, because of their bad literacy standards. So, I think it would be good to, erm, actually do those resits, but on the other hand, erm, I have nothing against a better teaching standard because there are a lot of unqualified teachers.

Liz   I think both is important, because, well, if you're not ready to go to secondary school, you'll always be left behind. That's not good. So, you should be tak- +taking, -- you should be given the chance to, erm, well, to be ready to go to secondary school, to lift your level of literacy and-and maths.

Dawn Yeah I think it is also very important, because, erm, children at that, er, age, erm, learn more quickly than later on. So, er, I think they don't have the chance anymore in secondary school to improve /their reading or math skills that much as in primary school because it's there to, erm, to create your basics.

Liz   /Exactly. Yeah, but what would help more? The better teachers or the, erm, resits? Because I think that both are very important, but I think that the solution they have it's-it's -- it sounds like it's either one or the other.

Dawn I don't think it should be either one or the other. I think it's necessary or important to, erm, to give those who are, erm, more behind extra time and this extra time can be given by more teaching or through resits. And with the resits they need more, erm, teaching, I guess. So if you combine both, that would be easier.

Liz   Yeah, I just think that it really sounded like you don't have to be a qualified teacher to teach. I think that's just... that's terrible.

Dawn I think they are just criticising David Cameron's, erm, unqualified teachers policy. I don't know to what extent they are unqualified, because they must have some education, but they want to introduce a new master teacher's status.

Liz   Mhm.

Dawn Erm.

Liz   I really don't know what that entails.

Dawn No, neither do I. But -- well. It sounds so fancy, I-I don't know if that would actually improve something. I think -- I really think that those resits are actually good because, erm, when you enter secondary school it is -- that is the point where you have to know all your basics and when you don't, you're already falling behind.

Liz   Yeah, well. If you compare it to Dutch, I think, erm, -- in that way our system is quite good. You know, after, erm, -- after primary school we're divided into groups, I'm not sure if I agree with that, but the primary school testing, I think that is good to see at what level you're at, and -- but if you're not ready for secondary school, you'll just repeat that year. /I think that- that is a very good system.

Dawn /I think that-. I think it also depends on, er, how many years you spend in primary school. Because I know that in the Netherlands they have, er, six years of primary school? Is that right?

Liz   Eight.

Dawn Eight. Eight years?!

Liz   Yeah, but the first two years are more like, like, erm.

Dawn Kindergarten?

Liz   Yeah.

Dawn Oh, okay.

Liz   Sorta.

Dawn Because in Germany you only have four years, and I think that's not enough /to.

Liz   /That's not enough.

Dawn No, really. I don't know how long they, er, spend-

Liz   -Er, I think seven.

Dawn Seven? Well, that's more than in Germany.

Liz   The er-.

Dawn   -Because I think what's really important is to give /them time.

Liz   /they keep repeating "of year seven."

Dawn   Alright.

Liz   So I think they have seven years.

Dawn   Yeah. Well..

Liz   Yeah.

Dawn   I think if you only do that in your last year, it would be sufficient.

Liz   Yeah, but-.

Dawn   -You shouldn't do it every year because then, as the critics say, then the tests would be too much and there is just too much focus on the tests (( )) instead of the quality of teaching.

Liz   But, erm, to fall back on the Dutch system again, erm. We have a practice test in the seventh year, erm, to see how you're doing. Erm, so then, erm, if you -- if you fail for this part, but do really well on that part, you don't have that pay much attention to that part and get more, erm, at- +attention, more /help for the other part.

Dawn   /attention for other parts?

Liz   Yeah, so that before you get the official test, you are already on the right level.

Dawn   Well, I think that is way better, because I don't remember doing any tests in Germany apart from the usual exams. But no real test that said, "Okay, now you are ready for, erm, higher education." And, well, it's really bad in Germany anyway.

Liz   Yeah, but I think the testing is a good idea but I think, erm, if they don't make it an official test, erm, but just, erm, an informal test to see how they're doing, and do that every year once and /see what to focus on.

Dawn   /I think that would be still formal.

Liz   Yeah, but, erm, instead of "you failed" it's just "now we can focus on this to help you further."

Dawn   Aah, just tests to keep track of the, erm, process.. /progress?

Liz   /Progress.

Dawn   Progress.

Liz   Yeah, so that before you have the official test and then fail miserably, you can get help beforehand.

Dawn   Yeah, I don't know if they actually already doing /that. 'Cause here they're only talking about the resit. I think, well, one disadvantage of the resit would be that you already get used to doing so many resits, like. Students, er, see -- I mean, even at university students often say, "Oh, I'm gonna do the resit, because I won't, er, succeed in this exam anyway." So, that might be a, erm, slight disadvantage.

Liz   /I don't know either. Yes, but on the other hand you could say that, erm, that a resit is, er, a way of -- well, it's a back-up.

Dawn   Yeah, it's a back-up.

Liz   Instead of "Oh, no, I failed this test. Now my life will be miserable because my secondary school, erm, experience will be always being left behind and things."

Dawn   Yeah, /I think something definitely necessary to -- before you go to secondary school.

Liz   /Erm, it's.. yeah, it's a back-up. Like, yeah, you're not lost forever because you failed this test. That's -- I think that's a very good thing. Erm.

Dawn   And better teachers would also help.

Liz   Yes, definitely. Yeah, I think, erm, if that many pupils are taught by unqualified teachers, I would not be surprised if most of those pupils fail that test.

Dawn   Or it -- wouldn't it be an option to give those pupils who are already falling behind some extra lessons? And -- I mean, you could do that instead of resits. Just some more tutoring for them, so they don't have to take the resit. But, I-I don't understand the critics' point. What it's -- why is it so horrible to do resits for primary school?

Liz   I don't agree with that either. And, you know, you can have a bad day and fail that test. And at least you have another chance.

Dawn   Well, they think, er, that this policy has confused teaching and testing and thinks that tests improve teaching. I don't think that's a valid argument, actually, /because-

Liz   /I agree with that. (( ))

Dawn   If you only do one final test at the end -- or after seven years, and you can still do a resit about that, it's not like you only focused on testing the pupils.

Liz   No, and tests don't improve teaching, tests prove how- how /well the teaching is.

Dawn   /they show how they are doing.

Liz   Yeah, that-that shows how the pupils are doing and how the teachers are doing as well. So...

Dawn   Well, it's very hard to determine who is actually to blame for the /bad test results if you -- if you have both unqualified teachers and bad or weaker students.

Liz   /Of course. Yes, but you can compare, you know, erm, -- in primary school it's just a class taught by one teacher. So, then you can compare that class to that class   and how badly one class did in that, er, subject and the other -- other subject and see if-if one subject is failed by more than half of the pupils in that class and very few pupils in the other class, then you know that something is wrong.

Appendix D: Transcription 3

Legend:

/	Said simultaneously
-	Interruption
--	Restart
+	Partial word
(( ))	Unintelligible speech

Pair Three

Jane   I have no idea where to start

ME   You have no idea where to start. What did you think of the article? Was it clear? Er, did you see anything that really caught your eye or anything? Or do you have a strong opinion on either side of the argument, 'cause, of course, there are two sides to the argument.

Lena   I can actually see both. I can understand that it would be -- a resit would be preferable. I mean, erm, because, everyone needs to have standards and everyone needs to be able to do math and English, especially living in an English-speaking country.

Jane   That's true, but that doesn't mean that you need standardised tests. I mean, those -- I think in Sweden they don't have standardised testings, they just, erm, I don't exactly know how their school system works, but they don't have standardised tests, erm, but they do have, it is also mentioned in the article here, erm, they have teachers who need to have a master's degree before they are allowed to teach. And that is apparently something that really helps, because er, the, er, the levels of education in Sweden are high.

Lena   Yeah, but-but it does help having qualified teachers.

Jane   Yeah, that's-that's really the problem I think. Because, I... Yeah, erm, more tests -- I'm not sure if that will help, as it -- yeah...

Lena   I think it might help if you don't have qualified teachers, but if you do have qualified teachers, er, the standard of teaching will be higher, so, you -- you won't need it as much. 'Cause now you have to see if the unqualified teachers are as good as the qualified teachers and if everyone (( )) the same. I -- yeah..

Jane   But I do think that these tests will probably just result in more stress for the students and more panic, like, "Oh, I need to, er, [laughs] I need to pass this test or else my future will be doomed."

Lena   I don't think that you think about it that way if you're in primary school. I mean, we have the CITO test -- I didn't think about them like that. I was, like, "Oh, it's just another one." And even the students, well, who were not that good at math or Dutch or whatever, they -- they were, "Oh it's just another test" not "my future is at stake, I can't (( ))". And more people had to resit year eight/. So, I think that it's a great idea if everyone has a start with the same standards in-in secondary school.

Jane   /That's true but --. Also with the CITO toets, erm, I had it as well and I wasn't worried either. But, generally I was a rather good student so got higher grades. So I wasn't worried, but I had classmates who were really worried and thought, like, "Oh, if I fail this one exam" or even if they, er, didn't get higher grades previously, they would, er, be worried that they had black-out during CITO toets. And then be worried, er, "Oh, shit, what-what if I get a low score and then aren't, erm -- I'm not able to, er, enter the school that I want to", because.. yeah. And, it is, of course, er, not just based on the CITO toets, er, score, er, where you, er, end up, beca- +because, -- it was a kind of, erm, advies, I think, advice, you get for er -- because it also based on how a teacher thinks of you, but with how big classrooms are, teachers are not always able to accurately, erm, describe, er, how good of a student someone is, because they don't have /time for all the students.

Lena   /No, exactly. We've got -- I actually got advice from my group four teacher who wrote notes for every student that he had in /class and he just kept them, so that he could -- when everyone's in group eight, it was, like, "Yeah! That was this student! He did really great!" And that really worked, because he noticed from everyone, well, we changed a bit, but not too much, 'cause then he noticed that "this one is an exceptionally good student." "This one needs a bit more encouraging, maybe, but then he can make it," and-.

Jane   /Oh, really? -But that's an exceptional teacher.

Lena   Yes, he is an exceptional teacher.

Jane   And not-not all teachers are, first of all, willing to do -- give that much effort.

Lena   He was a really young teacher, maybe that helped.

Jane   Yeah.

Lena   It was his first year of teaching actually, when I (( )).

Jane   That's kind of adorable.

Lena   Yeah.

ME   What if you were to look at tests in uni, for example? Do you have any particular opinions on that? '/Cause-.

Jane   /Fluency...

Lena   We talked about this yesterday, because, er, with fluency and, er, pronunciation you're basically -- you are, er, tested on how good you are now. And it's not compared to how good you were when you started. I mean -- I know I made a lot of progress from when I started here up till now. But still, if I pass my pronunciation with a six, I think -- it is barely good enough. And that really gets you down, while when I'm working I get compliments on how British I sound from a British native speaker. So, I think maybe you should also try compare it to where you started and just where you are now. Because -- I get it that there's sorta also a standard in there..

Jane   Yeah, that's kind of demotivating to get a six when you've worked very hard, yeah.

Lena   (( ))

Jane   And it's also -- er, I mean, I don't know how you are when you are in fluency tests, er --.

Lena   [laughs]

Jane   Yeah, I get really nervous. Er, suddenly, my vocab drops really, really low. 'Cause I repeat words, say 'like' and 'erm' so many times and just -- I get stuck in my own -- in my own mind a lot. And, er, and during my last, er, fluency exam I had to repeat myself so many times and I got stuck and had to say, "Sorry, I completely lost the train of thought," in the middle of a sentence because my mind just went blank /[laughs] because -- because of nerves. And if that's, er, that's the only moment you are gon- --gonna base the grade on and not just, er, how you're -- how you are doing in class -- that's kind of nerve wrecking.

Lena   /Oh, yeah. You get a sort of grade for what you do in class but that doesn't mean /that you passed or you failed. That doesn't really-.

Jane   /Yeah, that's -- that's -- that's the participation, er, thing, yeah.

Lena   I actually did quite well in this fluency test. Maybe because I talked with [student] this year and we can talk really well.

Jane   Yeah (( )).

Lena   Yeah, but you just get used to a certain person talking to you all the time.

Jane   Yeah..

Lena   And we also both had really weird topics, which helps too.

Jane   I did the exam with [student]. Well, I did it last year with her as well. And, well, we really get on and we have tons of discussions about any kind of topic. So, I would think that I would be prepared for that. But, I -- I was just, er, nervous about the teachers being there and the fact that I would -- was being graded on, because right now I can talk relatively fluently, because I'm not nervous about being graded and judged on my - [laughs].

ME   Good! Do you think we need more testing in general? Or do you think that testing just only puts more pressure on both the students and the teachers?

Jane   /I'm not completely sure.

Lena   /(( )) more testing. I think you should, er, -- instead of just having more /tests-.

Jane   /Like a gradual moments.

Lena   Yeah make it, indeed, more gradual, so that you have a sort of final exam to see if you can do it. But, then, you should also base that, the grade you get from that-that one exam, on what you actually did in class. In class -- in class I'm more fluent.

Lena   /So..

Jane   /Yeah.

Lena   And then you also contribute when you can, and I think that's also important. 'Cause now you're just forced to talk for ten minutes.

Jane   And also, during, erm, pronunciation lessons we didn't always get to talk that much. There was --

Lena   There were too many people.

Jane   Yeah, it's also -- how the course is set up and how much some teachers suck. I will not name names. We all know.

Lena   But I think it isn't -- that's not [the teacher]'s fault that the classes were that big. Because she actually had to make one group out of everyone and she managed to split them up. It was like, "I can't do anything anymore."

Jane   I mean this year we, er, got to talk more than we did last year. Because last year the groups were just ridiculously big.

Lena   Yeah, but I also think -- [the teacher]'s doing better this year. Because last year we gave her feedback, like, "If you failed pronunciation, please, give us some extra lessons and not just drop us and let us find it out on our own." Now she is actually teaching the other people who-who failed it. She -- she has taught them and it really does help. I had two extra classes with her and I passed. /So...

Jane   /That's good.

Lena   I don't know if it just me, or if she really helps, because she can do really good.

Jane   But yeah, it's just kind of weird for, erm, oral communciations class -- we don't do that much of talking. So seriously -- certainly not most -- more, erm, how do you say that? Spontaneous, er, speech. Because well we do talk in our lessons, of course, and we do, er, always speak in English, because [teacher] will (( )) you out, like, "English!"

Lena   Yeah, [teacher] does that too, but she often switch back -- switches back to Dutch herself, which makes it really confusing for us. 'Cause we're just all confident in English and, then, suddenly: Dutch. It's like -- like -- it's just...

Jane   Yeah, in my mind it would just make more sense for, er, OCS to have, erm, more spontaneous speech. Like, if we had -- discuss the news or something like every week that would already be more talking than, er --. And also those presentations, they take up so much time and I'm not sure exactly how much use they are.

Lena   Well, I think it -- it is of use. Maybe for the person who is doing talking then /and not for the rest of the class, 'cause the rest of the class is sitting there, like, "Oh, god. Not another presentation."

Jane   /Exactly. And for the one moment that's not, like, if you have that one class dedicated to you, basically -- that is not -- yeah, it -- it would be better --

Lena   Yeah, that's forty-five minutes for one person.

Jane   Yeah and [teacher] did mention at some point, I think, that, erm, er, in previous years, er, they would have the presentations, er, during EWOL, or whatever EWOL was before then. And that is was part of that course and that she would also be there to also judge the presentations and so then you would have more time during OCS for actual /talking.

Lena   /OCS.

Jane   So, I think that would be, er, a lot better already. Also, because the EWOL portfolio freaking sucked.

Lena   I think it was do-able.

Jane   It was do-able, but it's just the whole course was set up as -- no.

Lena   "Yeah, you have to buy a book and read the book." And I haven't done anything with it. It's a bit of a pity.

ME   Coming back to the article. For which camp are you? Do you think that the Tories have a good motion in issuing such a resit test?

Jane   I'm more inclined to go with Labour, er, like, just more qualified teachers. That helps a lot better, I think, than more standardised tests.

Lena   Yeah, I can agree with that, but if you can't get qualified teachers, then you should go for the testing, of course. /Just to see --.

Jane   /Maybe.

Lena   And you know where everyone is. I think that's also important. You need to know, when you secondary school, that, "Oh, this person needs some extra help in math" or "This person needs some extra help in English."

Jane   Yeah, but then you run the risk of what's also mentioned that kids are being, er, er, refused from schools based on their testing scores and that's not okay.

Lena   Yeah, but that happens here too. Because the school I was in was like, if you have lower than five-four-five on your CITO score /you can't enter.

Jane   /Really? Damn. But that's a really high score, though already.

Lena   Yeah, but -- gymnasium.

Jane   True, but still.

Lena   And we were a really small school. We had five hundred students when I started. So it's --.

Jane   I mean, what is the highest score you could get with CITO?

Lena   Five-fifty.

Jane   Yeah, exactly. That's (( )).

Lena   I think, like, erm, from the group of friends I still have from school, erm, I think, it's.. ninety per cent that actually had five-fifty and one person that just really fucked up and had five-forty-seven, but still got in, because it was like, "Yeah, he just has to get in."

Appendix E: Transcription 4
Legend:
/ Said simultaneously
- Interruption
-- Restart
+ Partial word
(( )) Unintelligible speech

Pair Four

Celia   So what do you think?

Anna   I think it's ridiculous. It's more adding tests for no reason. Because nothing happens, really happens when they fail.

Celia   No, because they just go on with the usual years they do it and. Yeah.

Anna   And then -- it just -- no. It puts a lot of pressure on the students. Blaming the students for the fact that they hadn't reached the level that the need to have. That's.. not how it's supposed to happen.

Anna   That's not very motivating.

Anna   No. And it not their fault. It's probably the teacher's fault. 'Cause it -- because it -- I had a number. That there are more than four hundred thousand pupils being taught by unqualified teachers. I think that is more the problem.

Celia   Yeah that's [laughs] you could say that is a slight problem, yeah.

Anna   Maybe. Just a little. That's not solved by simply adding a test.

Celia   No, because -- if-if the children fail, just testing it again won't make it any better.

Anna   No.

Celia   So, yeah.

Anna   Like I said, if they think that tests improve teaching -- That's not true.

Celia   That-that's not the case.

Anna   That is really not how it works.

Celia   Teaching should improve teaching and not the other way around.

Anna   Better education of teaching and -- I mean if they want the level of all, like, if... They want the level of every pupil to be the right level when they leave primary school, they have to change something about what they learn at primary school or the way they are taught at primary school or whatever.

Celia   Yeah, or just change the /standards.

Anna   /Yeah.

Celia   What they're supposed to know, 'cause not every child is good at learning.

Anna   Yeah. Exactly. And if they just add a test that doesn't really have consequences, it just puts the pressure a lot more on the pupil, and it's only they're, like, eleven at that point. And that's-

Celia   /And what's it gonna do?

Anna   /So, basically you're telling them already "Yeah, you suck. You fail. You can't -- you can't do it." So it just -- it doesn't motivate anyone.

Celia   Yeah, but what else is it gonna do? Like do a resit to prove that you've learnt something in the past months that you didn't know last year?

Anna   Yeah.

Celia   I mean, if-if you still don't -- do not pass any tests in the next year then maybe you should do that year again and not go on /and do the same tests over and over.

Anna   /Yeah, exactly.

Anna   It's just a red flag now and nothing is really done with it. So, basically they're saying that in the first year of secondary school teachers should be able to pull them up to the right level, but apparently schools are getting money for that.

Celia   Which I /didn't know.

Anna   /Which is sort of ridiculous because it's your job to get people to the level they're supposed to be at.

Celia   There shouldn't be extra money for doing /your job, which you're already paid for.

Anna   /Yeah, there's -- they are getting five hundred pounds for each student that doesn't have the expected standard after primary school, but maybe should -- they should put that money into improving in what they learn in primary school so that everyone meets that standard when they leave primary school.

Celia   Or improving those unqualified teachers.

Anna   Yeah.

Celia   And actually get good teachers in front of classrooms.

Anna   Yeah. So if you have someone that is really below the standard, keep them for a year.

ME   So you're really on the labour side then and that the Tories just..?

Anna   Yeah, they're blaming the student and that's ridiculous cause that is now how it works. Just -- it just -- it does not.

ME   If you gear this subject to university, do you think that universities -- do we have a focus on testing as well or /do you think there are-.

Anna   -Oh yes, definitely

Celia   Definitely.

Anna   Very, very result-focused.

Celia   It mostly depends on one or two tests you take per course and, well, if you happen to fail it... you're done. Even though you might have been doing well with the homework and all that during the whole course.

Anna   You can always have a bad day. It is very, very result-oriented.

ME   In your opinion, they don't need more testing, but do other /things-

Anna   /The teachers need more testing. The teaching -- the -- not necessarily the teachers themselves, but what is taught in class and how it is /taught.

Celia   /The whole teaching method.

Anna   Yeah, and to test -- to test if teachers are qualified.

Celia   And don't put unqualified teachers in front of children.

Anna   Because-because -- actually the way it is, I don't know how it is in the UK, but here it's -- the students are tested way more than teachers are, but I think that teachers should have a test every now and then to see if they are still qualified in doing their job. I mean, I do think that they that they do get, like, the courses in the between, like, if something changes. I think they're.. given courses on that.

Celia   Yeah. like mini workshops. /Something like that.

Anna   /Yeah. So, that it is kept up-to-date, they're updated. But if they use it, or actually do it that's the other question. But right now it's -- the focus is too much on just the student, I think. And a student can only do so much. When you're at university you're older, so you have a (( )) of your own responsibility. It shouldn't just be focused on the students.

ME   Yeah, do you think that in primary school it is right to have that many tests?

Anna   No, they are way too young.. for so many tests. If you introduce it that early, the stress levels will already ju- +just -- way too high when they reach pri- +primary -- secondary school and you are /creating..

Celia   /And it's going to get even higher by the time you get into college.

Anna   You're already creating a high pressure workload, pressure thing, for children. They're, like, not even ten for, like, through the biggest part of primary school. That's -- it doesn't work.

Celia   They shouldn't spend all afternoon reading books on history or whatever and just go play outside and do fun stuff.

Anna   Yeah, and if you introduce tests in primary school then (( )) also they have to start studying the way we do and that's, for that young age, is really not necessary. And it doesn't really work. They're too young, they're -- they don't have that level of high concentration, they don't see the importance yet, 'cause they're young.

ME   But, you know, we had CITO. How did you experience that, Celia?

Celia   I don't remember if I've actually studied for that.

Anna   No, I don't think anyone studied for that.

Celia   It was one test.. over the whole year. Well, not one test over the whole year, but one larger test over the whole year.

Anna   If that would've been the complete basis of which secondary school you would attend, then I'd would really object towards to it, /but it was a combination.

Celia   /It was more like a guideline.

Anna   It was more like a guideline and your own mentor was more the-the opinion you went on because he knew you and -- he knew you throughout the year. Because basing a-a big decision like that on purely on score of a test that you did on one day is not a really wise choice, but because it was more of a combination between the two then I think it's /..

Celia   /Yeah, because I know people who scored really low on the CITO, but were actually great students and they're -- they went to the Gymnasium and with straight A's, went to college. So, one test /doesn't..

Anna   /Maybe that is because it's not the test you study for /maybe that's

Celia   / (( )).

ME   So the pressure is the evil thing here, in your opinion.

Anna   Yeah, it is. And some tests are just not that necessary.

ME   So in this case, it really is not necessary to have a resit test.

Anna   Especially 'cause they don't do anything with it. The pupils aren't -- aren't held back a year or anything and they just use it as a red flag, but, that's it. There is /no..

Celia   /Yeah, I mean, if they actually had to do that year over again, then maybe it -- maybe you could use those tests, but if you let them go on-

Anna   -Yeah, cause indeed if they come on from primary school -- yeah, their level being to low for-for real secondary school and you find out through a test, then it might be good to

keep them in the first year of secondary school for a year longer. But, right now they're not really doing anything with the results.

Celia   No and it's seven tests, right?

Anna   Seven tests?!

Celia   Yeah, seven resit tests. Why would you do seven resit tests?

Anna   Yeah, 'cause if you do that, that even increases the feeling of failure inside the student -- inside a pupil, like, "I have to do it again."

Celia   "Again, and again, and again. Next month I have to do the next one again."

Anna   For a eleven year old that's really heavy stuff.

Celia   I wouldn't be happy if I had to take all those resits.

ME   Do you have any other things that have caught your attention?

Anna   I think it's wonderful that they, like, blame Labour for that, erm, one in three children left primary -- primary school unable to read when it's not really the pol- +politics, like, how can it really be their fault when it's more inside of the education -- that -- /I think --

Celia   /it's an educational system.

Anna   I think they blame the whole, erm, thing in politics for something that goes wrong on a different level that doesn't really -- It's -- It's great -- it's brilliant that they blame each other for things that they can't blame each other completely for, and they still do it. But, people --/.

Celia   /Well, that's politics.

Anna   Yeah, and people believe it, so that it works. But it's really funny to see how they do that.

ME   Yeah, they are really juxtaposed in this article. Exactly.

Anna   And I don't -- I also had the idea it is not that bad in the Netherlands the way they do that than in Britain. Maybe that's because we have more parties or something like that.

Appendix F: Permission Sheet

Dear participant,

First of all I would like to thank you for being here today! I greatly appreciate it that you are willing to help me and my research. It all brings me one step closer to finishing my thesis.

Quick overview of what we are going to do:

- You will discuss the article with your partner. Read it thoroughly.

- You may make some notes on the article as to what you want to discuss or things that caught your attention.

- Feel free to think beyond the article! You are allowed to discuss other relevant themes that are relevant to this piece of writing.

- The talk will last about 5 to 10 minutes.

Since I want to capture your speech in its most spontaneous form, I would like to ask of you NOT to talk about the article and/or research with anyone after you have left this room. You might unconsciously influence a fellow participant, their data and inherently my research.

Have fun and, again, thank you for participating!

I hereby give the researcher permission to make (video) recordings during the experiment and declare that data gathered during the experiment may be used for research purposes.

_____

Name

_____

Signature