

Department of Modern Languages  
University of Helsinki  
Helsinki

**ENGLISH-MEDIUM INSTRUCTION: SEEKING  
ASSESSMENT CRITERIA FOR SPOKEN  
PROFESSIONAL ENGLISH**

Diane Pilkinton-Pihko

ACADEMIC DISSERTATION

To be publicly discussed, by due permission of the Faculty of Arts  
at the University of Helsinki in lecture room 10, University main building,  
on the 13<sup>th</sup> of December 2013, at 12 noon.

©Diane Pilkinton-Pihko 2013

ISBN 978-952-10-9519-1 (paperback)

ISBN 978-952-10-9520-7 (PDF)

Unigrafia

Helsinki 2013

## Abstract

In higher education, the desire to internationalize has created demands for an internationalized academia to use English increasingly in teaching outside the English native-speaking world. Given this situation, perhaps other criteria for measuring successful communication should be considered than that of the native-speaking minority. With lecturers whose native language is not English increasingly teaching their subjects through English, there is a growing need to develop adequate measures for this purpose and situation as the current normative standards are no longer tenable. Establishing adequate measures for this purpose and situation are relevant to institutions facing the challenge of providing EMI courses and programs while ensuring credible quality control.

In order to determine what criteria might be adequate for assessing spoken professional English in an international context, this study investigates self-assessments of professional language in relation to language ideologies. The study involves English-medium instruction (EMI) in the field of engineering and takes place at a Finnish university. Using a mixed-methods approach, the study employed an explorative strategy that involved a concurrent design. The two methods were used in parallel and the results integrated at the interpretation phase. This approach provides a general picture through micro- and macro-level analyses: the self-perceptions of EMI lecturers (i.e. qualitative) and their students' perceptions of English in lectures (i.e. quantitative). The investigation employs a bottom-up approach, and is primarily qualitative. The findings are based on authentic data: video-recorded interviews and lectures, their transcriptions, and a questionnaire.

The findings show that EMI lecturers have two basic representations of their English: A) when they compare their English to native-like targets, they find fault with their English, and B) when they think of themselves in their normal work environment, they see their English as working rather well. Certain language ideologies induced type A discourse, including standard language and NS language ideologies, and others induced type B discourse, such as English-as-a-global-language ideologies. The results from the student questionnaire also support interpretation B.

Since meaningful testing should reflect the target situation, what my informants say in the type B discourse is relevant to developing assessment criteria. Their views to Common European Framework of Reference for Languages (CEFR) scales are also extremely useful in pointing the way towards the central elements upon which relevant assessments for professional English in an international environment should be based. The conclusions indicate a comprehensibility goal over native-likeness for assessing spoken professional English in an international context. The study outlines some criteria relevant for assessing spoken English for this purpose and situation.

## Acknowledgements

Writing this dissertation has been quite a journey. The inspiration initially began with a work-related project involving development of a spoken skills test for engineering students at Aalto University, and then continued with another work-related project involving mentoring and assessing academic teaching staff in EMI programs. In both cases, English as a lingua franca and assessment were central themes. In the end, I pursued studying these themes from the perspective of EMI lecturers. In the early stages of my studies, part of my journey (in life) also took me to India for two years. I am convinced that living in what felt like total chaos for two years taught me new skills important for managing day-to-day life, which have also been invaluable for completing this dissertation.

Along the way, I have met and received help from several people, all of whom contributed in various ways. Without their support and encouragement, it would have been impossible to complete this study. Now, I would like to extend my sincerest gratitude to all those who came to the rescue when I needed it and who played a central role in the successful completion of this dissertation.

First, I express my deepest gratitude to my doctoral supervisor, Professor Anna Mauranen, whose brilliance was instrumental in shaping my work. As an outstanding researcher and visionary, she is an exemplary figure on many fronts. I also thank her for the excellent feedback and consistent encouragement.

I am also grateful to the preliminary examiners, Professor (emeritus) Sauli Takala (University of Jyväskylä) and Dr Alessia Cogo (University of Southampton) for their insightful comments and constructive feedback, which helped me to consider my work from the perspective of potential readers in my professional field.

To the lecturers who participated in this exploratory study, I thank you for sharing your valuable time and for being flexible. Your willingness and courage made this study possible, allowing data to be collected through interviews, observations, and video sessions. I also thank your students for their willingness to participate and for completing the questionnaires.

I am especially grateful to Professor Sebastian Pannasch, whose guidance and assistance in the statistical analysis undertaken was strategic. I also thank Dr Matti Taavitsainen for helping me in getting started with my statistical analyses.

A special thanks also goes to the SELF (Studying in English as a Lingua Franca) group, whose camaraderie at seminars, conferences, and other events has been an important source of inspiration. As the members of this group have changed over the years, I extend special thanks here to those with whom I have had the most contact: Ray Carey, Henrik Hakala, Niina Hynninen, Elina Ranta, Jaana Suviniitty, and Svetlana Vechinnikova. I also extend my warmest thanks to Niina Hynninen and Jaana Suviniitty for sharing their data collection instruments with me, and to both for being loyal and caring friends. To Jaana, I also thank you for having made this experience fun at opportune times.

I am also grateful to colleagues who supported and encouraged me throughout the process of studying, writing, and rewriting. A special thanks goes to Dr Heidi Rontu for

the enthusiastic discussions in recent years and whose encouragement helped me to finish sooner than I might otherwise have. I also owe a special thanks to my colleague, Laura Mendoza, for her understanding and continuous support during the writing of this dissertation. To my other colleagues, I also extend my gratitude to those who have given encouragement and support during my studies.

Closer to home, I express gratitude to my family, especially to my husband, Pete Pihko, whose encouragement and understanding gives me strength, and to our children, Jekaterina and Vilho, who have helped to keep me present in everyday life. My heartfelt thanks extend to all of you for lending support in both emotional and practical matters.

# Contents

<b>Abstract</b>	3
<b>Acknowledgements</b>	4
<b>Abbreviations</b>	13
<b>1 Introduction</b>	15
1.1 Language ideologies informing L2 performance assessment	16
1.2 English-medium instruction	19
1.3 Research objectives	21
1.4 Research questions	21
1.5 Overall research design	23
1.5.1 Research methodology	23
1.5.2 Theoretical framework	24
1.6 Overview of the dissertation	25
<b>2 Internationalization</b>	27
2.1 Background to the study	27
2.2 Why focus on English-medium instruction?	28
2.3 European education: language policies and language ideologies	31
2.3.1 Higher education in Finland: Language policies	32
2.3.2 Aalto University: Language guidelines	33
2.3.3 From language policy to institutional practices: two examples	34
2.4 International communication in English	37
2.4.1 What is English as a lingua franca?	37
2.4.2 NS ownership controversy	38
2.4.3 Idealized NS or competent SLU?	39
2.4.4 SLU vs SLA	40

2.4.5	Some key findings in ELF research	41
2.5	Summary	43
<b>3</b>	<b>Language ideologies and their social position</b>	<b>45</b>
3.1	The notion of language ideologies	45
3.2	Language ideologies interfacing with social action and structure	49
3.3	Language ideologies as gatekeepers	50
3.4	My definition of language ideologies	51
3.5	Some widespread language ideologies	52
3.5.1	One-nation-one-language ideologies	53
3.5.2	Standard language ideologies	53
3.5.3	Native speaker language ideologies	56
3.5.4	English-as-the-global-language ideologies	59
3.6	Conclusion	60
<b>4</b>	<b>Assessing spoken English: ELF, LSP, and self-assessment</b>	<b>61</b>
4.1	Performance assessment	61
4.1.1	Performance in second language assessment: two traditions	63
4.1.2	Performance-Based Assessment Criteria	65
4.2	Spoken ELF in assessing English performance	69
4.3	Assessing language for specific purposes (LSP)	75
4.4	Self-assessment in professional assessment	77
4.5	Chapter summary	79
<b>5</b>	<b>Reviewing the ideological basis of an assessment tool: CEFR</b>	<b>80</b>
5.1	CEFR – an overview	80
5.2	CEFR - Language description versus prescription	84
5.3	Examining the CEFR for language ideologies	86
5.3.1	Linguistic competence	87

5.3.1.1 On grammatical competence	88
5.3.1.2 On lexical competence	89
5.3.1.3 On phonological competence	93
5.3.2 Sociolinguistic competence	95
5.3.3 Pragmatic competence	98
5.4 Conclusion	102
<b>6 Research design and methods</b>	<b>103</b>
6.1 Rationale for the mixed methods approach	103
6.2 Participants	106
6.2.1 Overview of primary and secondary participants	105
6.2.2 Primary participants	106
6.2.2.1 Lecturer A: L1 Arabic and French	106
6.2.2.2 Lecturer B: L1 Finnish	107
6.2.2.3 Lecturer C: L1 Finnish	107
6.2.2.4 Lecturer D: L1 Dutch	108
6.2.3 Students	109
6.3 Sources of data	109
6.3.1 Triangulation strategies	110
6.3.2 Interviews	111
6.3.3 Stimulated recall	113
6.3.4 Repeated survey	115
6.3.5 Other methods	116
6.4 Analysis of the data	117
6.4.1 Qualitative data analysis	117
6.4.2 Quantitative data analysis	118
6.5 Validity and Reliability	119



<b>7</b>	<b>Analysis of language ideologies</b>	121
7.1	One-nation-one-language ideologies	121
7.2	Standard language ideologies	123
7.3	NS language ideologies	129
7.3.1	A hierarchy of accents	129
7.3.2	NS – idealized vs real	132
7.3.3	Different sociolinguistic appropriateness for different social contexts	134
7.4	English as the global language ideologies	136
7.5	Conclusion	138
<b>8</b>	<b>Analysis of self-assessments in relation to five CEFR scales</b>	141
8.1	CEFR self-assessed English	141
8.2	Sufficient CEFR levels for lecturing engineering subjects	142
8.2.1	Telling a story logically is important and jumpiness is normal	143
8.2.2	Accessible (NNS) fluency is central	145
8.2.3	A reasonable accent is fair	148
8.2.4	Non-native features are acceptable	152
8.2.5	Specialized vocabulary is primary	155
8.3	CEFR limitations	158
8.3.1	Limitations for LSP	158
8.3.2	Limitations for ELF	159
8.4	Conclusion	160
<b>9</b>	<b>Analysis of students' perceptions of English in lectures</b>	163
9.1	Results	163
9.1.1	Likert items	163
9.1.2	Multiple-choice items	167
9.1.3	Open-ended item	168

9.2 Discussion	171
9.3 Summary	174
<b>10 Conclusions and implications</b>	<b>176</b>
10.1 What language ideologies guide EMI lecturers' self-assessments of their professional English?	176
10.2 To what extent are the five CEFR scales a potential tool for assessing the professional English of EMI lecturers?	179
10.3 What are EMI-students' perceptions of their lecturer's English during lectures, and do their perceptions change during the course?	181
10.4 What qualities of spoken professional English in an international context are relevant to developing assessment criteria? Implications	183
10.5 Evaluation of the study	187
10.6 The future	189
10.6.1 Taking a step forward	189
10.6.2 Future research	190
<b>References</b>	<b>192</b>
<b>Appendix A: Self-analysis</b>	<b>213</b>
<b>Appendix B: Guide for interview 1</b>	<b>215</b>
<b>Appendix C: Guide for interview 2</b>	<b>216</b>
<b>Appendix D: Guide for interview 3</b>	<b>217</b>
<b>Appendix E: Guide for stimulated recall</b>	<b>220</b>
<b>Appendix F: Student questionnaire</b>	<b>221</b>
<b>Appendix H: Transcription conventions</b>	<b>223</b>

## List of Tables

TABLE 1-1 KEY THEORETICAL CONCEPTS THAT APPLY TO THE PRESENT STUDY	25
TABLE 4-1 FACETS OF VALIDITY FROM MESSICK 1989	67
TABLE 4-2 WHAT QUALITIES MEASURE PROFICIENCY OF SLLS AND SLUS?	71
TABLE 5-1 CEFR SCALE OF GRAMMATICAL ACCURACY	88
TABLE 5-2 CEFR SCALE OF VOCABULARY RANGE	91
TABLE 5-3 CEFR SCALE OF VOCABULARY CONTROL	92
TABLE 5-4 CEFR SCALE OF PHONOLOGICAL CONTROL	94
TABLE 5-5 CEFR SCALE OF SOCIOLINGUISTIC APPROPRIATENESS	97
TABLE 5-6 CEFR SCALE OF COHERENCE AND COHESION	99
TABLE 5-7 CEFR SCALE OF SPOKEN FLUENCY	99
TABLE 6-1 OVERVIEW OF PRIMARY (A-D) AND SECONDARY (E-G) PARTICIPANTS	105
TABLE 6-2 SELF-ASSESSED ENGLISH FOR PROFESSIONAL AND NON-PROFESSIONAL PURPOSES	106
TABLE 6-3 OVERVIEW OF LECTURERS (A-D) AND THEIR STUDENTS	109
TABLE 6-4 DATA AND METHOD TRIANGULATION STRATEGIES	110
TABLE 6-5 EXAMPLES OF FUNNELING TECHNIQUE	113
TABLE 8-1 SELF-ASSESSED ENGLISH ON FIVE CEFR SCALES FOR PROFESSIONAL AND NON-PROFESSIONAL PURPOSES	142
TABLE 8-2 FIVE CEFR SCALES: LEVELS LECTURERS CONSIDER SUFFICIENT FOR LECTURING ENGINEERING SUBJECTS	143
TABLE 8-3 SOME CEFR LIMITATIONS FOR ASSESSING LSP	158
TABLE 8-4 SOME ASPECTS TO CONSIDER FOR ASSESSING LECTURING	159
TABLE 8-5 SOME SUITABLE CRITERIA FOR ELF - MISSING FROM CEFR	160
TABLE 9-1 ANOVA RESULTS OF WITHIN-SUBJECTS CONTRASTS	164
TABLE 9-2 ANOVA TWO-FACTORIAL REPEATED MEASURES RESULTS FOR TESTS OF BETWEEN-SUBJECTS EFFECTS	165
TABLE 9-3 CORRELATION BETWEEN PERCEIVED LECTURE COMPREHENSION AND PERCEIVED LECTURER ENGLISH AT TIME 1 AND TIME 2	166
TABLE 9-4 STUDENTS' COMMENTS ON LECTURER DELIVERY SKILLS	168
TABLE 9-5 STUDENTS' COMMENTS ON DELIVERY OF LECTURE CONTENT.	169
TABLE 9-6 STUDENTS' COMMENTS RELATED TO LECTURE COMPREHENSIBILITY	169
TABLE 9-7 STUDENTS' COMMENTS RELATED TO PRONUNCIATION	170
TABLE 9-8 STUDENTS' COMMENTS REFLECTING ELF VIEWS	170
TABLE 10-1 FINDINGS – SOME RELEVANT CRITERIA FOR ASSESSING SPOKEN PROFESSIONAL ENGLISH IN AN INTERNATIONAL CONTEXT	184
TABLE 10-2 FINDINGS – SOME RELEVANT CRITERIA FOR ASSESSING LECTURING	185

# List of Figures

FIGURE 1-1 VISUAL REPRESENTATION OF THE STRUCTURE OF THIS STUDY	26
FIGURE 4-1 CONTINUUM OF LANGUAGE CRITERIA IN PERFORMANCE ASSESSMENT, BASED ON MCNAMARA 1996	66
FIGURE 4-2 ROUGH CHECKLIST FOR DETERMINING TEST-TAKER NEEDS	74
FIGURE 9-1 QUESTIONNAIRE MEAN RESPONSE AT TIME 1 AND TIME 2 FOR EACH THEME	165
FIGURE 9-2 SCATTER PLOT SHOWING MEAN VALUES AT TIME 1 AND TIME 2 AS MEASURED ON A 4-POINT LIKERT SCALE	166
FIGURE 9-3 Q33 – STUDENTS’ PERCEIVED ABILITY TO FOLLOW THE LECTURER’S ENGLISH AT TIME 1 AND TIME 2	167
FIGURE 9-4 Q34 – WHAT STUDENTS’ PERCEIVED AS CHALLENGING IN THE LECTURER’S ENGLISH AT TIME 1 AND TIME 2	167

## Abbreviations

BELF	Business English Lingua Franca
CEFR	Common European Framework of Reference for Languages
EAP	English for academic purposes
ELF	English as a lingua franca
ENL	English native language
ELT	English language teaching
EMI	English-medium instruction
ESP	English for specific purposes
L1	First language
L2	Second language
LFC	Lingua franca core
LSP	Language for specific purposes
NS	Native speaker
NNS	Non-native speaker
SLA	Second language acquisition
SLL	Second language learning/learner
SLU	Second language use/user



# 1 Introduction

In a flat world of global players, English is the lingua franca (ELF) for many professionals, which includes a wide array of domains: from politics to entertainment, news media to social media, and academic lecturers to hairdressers – to mention a few. With increased mobility and communication, our global village operates in English, where the number of non-native speakers represents the majority (approximately 74% of face-to-face communication; Graddol 2006: 29). The spread of English has been and still is being enhanced by what Thomas Friedman (2005) calls the flattening of the world, a world in which borders are less visible and new technologies are exploding. With new global forces at work, English has become the language of choice for international and multinational companies, a primary example of language for specific purposes (LSP). With its widespread use in English-medium instruction (EMI), it has also become the language of choice for academia (known as English for academic purposes or EAP), especially in higher education – although there may be few ENL speakers present. In this international context, ELF users shape English to their own purposes and this occurs on different linguistic levels (see, for example, Firth 1996; House 1999; Mauranen 2006a on pragmatics; Jenkins 2000 on phonology; Mauranen 2012 on lexis/lexicogrammatical features).

The difference in usage between ELF and English native language (ENL) has given rise to debates in applied linguistics about norms and standards, which includes questioning the ideological basis of English language teaching and assessment (see, for example, Jenkins 2000, 2006a, 2006b, 2007; Taylor 2006). Both the debates and the international context in which I work gave impetus for the present study. My work at Aalto University in a Language Centre led me to mentor university lecturers who were providing English-medium instruction (EMI). As part of a pilot-mentoring program, these EMI lecturers could also optionally participate in a pilot-certification assessment for teaching in EMI (for details, see 2.1). Against this backdrop, I began to question the adequacy of ENL-based assessment criteria for the evaluation of professional English in this international context. Thus, the central question in this study explores what qualities of professional English in an international context are relevant to the development of adequate assessment criteria for this purpose and situation.

To achieve this, I investigate the self-assessments of professional English by EMI lecturers. These professionals use English to disseminate knowledge in classrooms where few, if any, ENL speakers are present. This situation stems from the desire of universities to internationalize, which creates a major motivation for them to offer increasingly EMI programs, especially at the master's level. This is the case in Finland as well as in many other countries around the world where English is not the local language, nor the native language of the majority of EMI lecturers or students. Being a player in this international market puts pressure on universities to provide high-quality education in English, thus universities are moving towards implementing measures for assessing the quality of EMI. These circumstances have led to a change in the job demands of lecturers, who now need to teach through English and at a level that meets the quality assurance goals set by the university.

My focus on spoken professional English includes data from EMI lecturers and their students. The study is carried out at a Finnish university and deals with lecturers of engineering. The international context provides a natural setting for investigating professional English that requires orally communicating complex ideas on demanding topics. Since its speakers are primarily second language users (SLUs, cf. Mauranen 2011b), an ability to communicate through English also relies on flexible language use that heightens variability. Because EMI lecturers are educating the future generations through a lingua franca, they are an influential group that could have a normative influence (Mauranen 2011a). Moreover, given that ENL speakers have a minority position in a global English-speaking village, their concept of what constitutes ‘good English’ may not be a valid view. These complex circumstances provide interesting grounds for the study.

This chapter first situates the study in the general framework of language ideologies and L2 performance assessment, followed by an overview of the general context of EMI, after which it presents the research objectives and questions. The overall research design is then explained and the chapter finishes with an overview of the structure of this dissertation.

## **1.1 Language ideologies informing L2 performance assessment**

Language ideologies provide a channel for understanding notions about language and for analyzing their relation to social phenomena. As a research instrument, they open possibilities to study perceptions about language as connected to social context. As they inform an individual’s language use in terms of what is viewed as sociolinguistically appropriate or correct, they can provide fruitful information into views held about assessing language for a given purpose and situation. It is this function of social evaluation that places language ideologies at the center of language assessment.

While language ideologies are important to assessment, so is social context. This latter aspect, however, has not yet been embraced or developed within a theory of L2 assessment (Baumgardner and Brown 2003; McNamara and Roever 2006). These scholars connect its disregard to ideologies related to monolingual NS norms. Jenkins (2007) further argues that a source of conflict for assessing ELF users stems from current assessment criteria being rooted in standard language and native speaker (NS) language ideologies. Moreover, as the literature rightly points out, L2 speakers are entitled to choice in terms of appropriateness for context (e.g. Jenkins 2000, 2007; Seidlhofer 2001, 2004, 2011; Widdowson 1994). What the literature suggests is that the needs of the local context should inform appropriateness, not arguments of NS superiority or ownership (ibid). In L2 assessment, this means that L2 speakers are directly involved in choosing their point of reference as either ELF or ENL according to their perceived needs. Against this backdrop, the present study approaches L2 performance assessment from a language-ideological perspective, by exploring perceptions of professional English in its local context.

The concept of developing alternative assessment measures for ELF users, however, is not all embracing. Attitudinal research indicates that the vast majority of ELT



professionals, some linguists, and reports in news media reject the notion of ELF. A recent study shows that non-ELF views are based on pragmatic reasoning, not empirical research (Jenkins 2007). For example, most ELT professionals adhere to NS language ideologies, which is not surprising since they formed the basis of their training at school as well as their own personal targets (ibid). Similarly, some linguists still view one standard as sufficient, such as Quirk (1985; 1990), who advocates Standard English as the best for all users regardless of the context of usage. Explicit in his argument is the belief in standard language ideologies.

A key ingredient in standard language ideologies is uniformity of language. As Deborah Cameron points out, this ideology presumes that "variation is deviant; and that any residual variation in standard English must therefore be the contingent and deplorable result of some users' carelessness, idleness or incompetence" (Cameron 1995: 39). This belief entails that each variety of Standard English, such as British or American English, is uniform and that any aberration from it is unthinkable. Thus, the language produced by learners must conform to the uniform standard. In accordance, the job of ELT teachers is to teach students a variety of Standard English. Many tools of the ELT profession are also instruments of standard language ideology, such as grammar books, dictionaries, textbooks, teaching manuals, and standardized language tests, all of which give the illusion of a uniform standard language, referred to as the 'target language' in ELT contexts. Probably one of the most striking indications of standard language ideology is the treatment of 'error' in language teaching, where deviations from native-likeness are referred to as 'L1 interference', 'interlanguage', or 'fossilization' (see Selinker 1972, 1992). Rating scales against which performance is judged also include wording that signals this deficit view by making reference to L1 features or 'foreign' accent, for example.

Other attitudinal studies have also investigated views held toward different varieties of English. Studies involving ELT practitioners and teacher trainees (e.g. Majanen 2008; Sallinen 2009; Seidlhofer and Widdowson 2003) as well as SLLs and English teachers (e.g. Hakala 2007; Pihko 2004; Ranta 2004; Ranta 2010) include views on what English is preferred and what English should be taught to learners. What the findings imply is a preference for ENL varieties as the model to be taught to SLLs, although attitudes appear to be opening up to accepting other varieties. One study also implied that ELT professionals held different views towards English outside of the language classroom (Majanen 2008). This study indicates that ELT professionals recognize that lingua franca communication is well tolerated in the real world (ibid: 64). Lending support to this view are several studies on professionals working in ELF contexts, including large-scale studies on EMI lecturers who generally report their English to be sufficient (e.g. Jensen et al. 2009; Pilkinton-Pihko 2011; Vinke et al. 1998) and studies on business professionals communicating through English as the language of global interaction (e.g. Ehrenreich 2011; Louhiala-Salminen and Kankaanranta 2011). While the views of ELF users show tolerance for variety in English, it contrasts with views reported in the media. For example, a study in Denmark was undertaken on academic ELF when it was portrayed as "Circus English" in the news media (Mortensen 2008). In Finland, the public views are also present in the news media, as indicated by the lack of tolerance for the non-ENL

accents of public officials (Nousiainen 2009). What these attitudinal studies and opinions in mass media provide is a glimpse of the preferences, the majority of which signal standard language ideologies. A point to bear in mind in relation to the attitudinal studies is that the views of ELT teachers, SLLs, and the news media represent speakers who do not regularly engage in ELF communication themselves, and thus reveal very little about the experience of using ELF.

With the rise of ELF, a renewed debate on language norms and the nature of ENL proficiency measures emerged. In a state-of-the-art article on testing, Leung and Lewkowicz (2006) criticize ENL “self-imposed normative” assessment practices as untenable (ibid: 28). Other research also raises concerns about the nature of current ELT assessment, including McNamara (2005) on ‘shibboleth-like tests’, McNamara and Roever (2006) on the social dimension in language testing, Shohamy (2001) on the issue of power, and Kim (2006) on the inappropriacy of ENL standards for international contexts. A recommendation for testing ELF was also proposed in Jenkins (2006a). She suggested a shift away from ENL correctness towards a pluricentric approach that focuses on successful communication in ELF interaction (ibid: 49). The response from Cambridge-ESOL explained their testing policy and practice, and gave no indication of making changes (Taylor 2006). Such a response positions their view with that of Quirk and other supporters of standard language ideologies. It also highlights ENL normative values that are irrelevant for ELF, a contact language that has no definable ENL target culture. Rather the target audience is international or global. This debate over language norms and proficiency measures for English centers on three language ideologies: standard language and NS language ideologies versus English-as-a-global-language ideologies.

In contrast to an emphasis on correctness, ELF users may be unconcerned with this aspect of language as illustrated by the use of regularized verbs and nominal plurals. While this usage may suggest a lack of accuracy, it does not necessarily mean that it causes comprehensibility problems. In some cases, the inaccuracy may also equate ENL usage but be deemed unacceptable for L2 speakers. One example is subject-verb agreement in existential sentences such as, ‘there’s three cats on my car’, which is common in spoken language among standard ENL speakers. Findings from Ranta (2009) show *there’s + plural noun* to be more common in an ENL spoken corpus (i.e. MICASE) than in an ELF spoken corpus of academic language (i.e. ELFA). Even though in common usage, this item is considered grammatically incorrect when produced by L2 speakers. Moreover, test scores based on such prescriptive ENL values are treated as objective realities, viewed as natural and representative of the real world. These ‘objective realities’ also inform the decision for acceptance to a desired position, as determined by the score from a high-stakes language test. For instance, L2 speakers planning to study in an international program taught in English at the tertiary level are generally required to pass a standardized language test, such as TOFEL, IELTS, CPE, or other similar tests. Such assessments, based on normative values and prescriptivism, may be useful for studying in ENL countries but questionable outside the ENL world.

The difference in context of usage for ENL and ELF poses new demands for assessing L2 performance, where high levels of language proficiency cannot be defined by native-

likeness for international contexts. This situation indicates a need to develop a new concept free of nativeness, such as the ‘educated speaker of English’ (Mauranen 2012: 238), which includes features of a ‘good communicator’ in a non-ENL matrix (ibid: 239). The ELF literature also points toward re-conceptualizing phonological, lexicogrammatical, and pragmatic competences (e.g. House 1999; Jenkins 2000; Mauranen 2012; Seidlhofer 2004, 2011). Re-conceptualization includes the ability of a plurilingua-cultural speaker to utilize his/her acquired linguistic repertoire (i.e. knowledge and skills) while maneuvering between different varieties of English (Canagarajah 2005) and skillfully adjusting to the linguistic demands present in an international context.

Reviewing attitudinal research provides an understanding of preferences, and signals a need to reach beyond them. For instance, studies that represent views of individuals who are not ELF users themselves do not tell us much about the views of those who actually use ELF. This includes ELT practitioners or trainees who may need to consider whether ELF will become an aspect of their teaching, but who may not have practical experience in using ELF. For non-ELF users who lack experience communicating in ELF, it may be difficult to envision what ELF entails or means (Hynninen 2010: 30). Similarly, views in mass media appear to represent voices that fear domain loss in national languages as well as fear of language change or losing standards. These voices contrast with those who actually use ELF for communication in their work. To advance beyond ELF attitudinal research and debates over proficiency measures, it would be important to explore how ELF users talk about their English in relation to its context of use. As regular users of professional English in an international context, their views could provide meaningful insights into what good communication in a lingua franca entails. Such information is central to the development of meaningful criteria for assessing English for this purpose and situation.

## **1.2 English-medium instruction**

This study was conducted in the context of EMI, where the lecturers were receiving training for teaching in English in a multicultural environment (for details, see section 2.1). As a mentor in the program, I provided guidance to four EMI lecturers. This guidance included giving feedback on the implementation of input received in the mentoring workshops as well as discussing any other relevant aspects of their teaching raised by them. (For an overview of the research design, see section 1.5.) This situation provided a natural environment for collecting the data for this study.

Because the data were collected from EMI teaching events, the speakers represent SLUs of English, not SLLs. In this context, they are professionals using English to accomplish their work. Having this role, however, does not mean that they cannot assume other roles. For example, in the mentoring program, they are also learners of how to teach effectively in English in a multicultural environment. In this role, they may refer to their English as being “too good” or “good enough”. The situation itself does not place them in the position of language learner. The distinction between ‘learner’ and ‘user’ is important

since these two roles belong to different research paradigms, the former to second language acquisition (SLA) and the latter to SLU (see discussion in 4.2).

As professionals, ELF users have different goals and needs than SLLs in a language classroom. Their jobs demand that they perform work-related tasks to a professional level in a multicultural environment. The knowledge and skills required for this purpose and situation differ from those required for native-like proficiency that is contrasted against ENL ‘target cultures’ as models of sociolinguistic appropriateness (e.g. Jenkins 2000, 2007; Mauranen 2005a, 2007; Seidlhofer 2001, 2004, 2011). In an international context where English is the language chosen out of necessity, the situation is quite different in terms of communicative demands and needs (Mauranen 2011b). In this context, communication hinges on an ability to work within various cultural mixes as they arise (ibid). The conventions necessary for working with this variability can hardly be linked to ENL sociolinguistic appropriateness (ibid). Nor can ELF as an ‘instrument language’ (cf. Mauranen 2011b) be confused with a ‘target language’. These basic distinctions also signal the inappropriateness of ENL-targeted L2 performance assessments as measures of professional English for international contexts.

On using professional English, ELF research shows that speakers take a ‘user’ view to English, not a ‘learner’ view. Studies on perceptions of using ELF to accomplish work-related tasks imply a goal orientation, where using English is described as “simply work” (Kankaanranta and Louhiala-Salminen 2010: 207; see also Kankaanranta and Planken 2010) or as signaling work-related goals: “it isn’t the purpose of a lecturer to show that he is good in English” (Pilkinton-Pihko 2010: 68). Findings from Ehrenreich (2009) also show speakers to be focused on conducting business, where members accept endonormative language use. These orientations point toward interest in communicating efficiently within a community of practice (cf. Lave and Wenger 1991), where correct language is not the central aim.

Since the late 1990s, internationalization has become a common strategy in universities and on a global scale (discussed in Chapter 2). With it came an explosion of EMI programs. The clientele consists of a broad network of professionals, including students, educators, trainers, and administrators. As ELF has become mainstream, particularly in the global world of business and education, the work environment in higher education now resembles that of an international corporate office. Even though this study is conducted with EMI professionals, their international context overlaps with professionals in other fields. The findings will therefore be of professional interest to those working with speakers using professional English in international contexts, and this does not exclude ENL speakers – who also need to demonstrate skills necessary for working in an international context – where being a monolingual ENL speaker could be a disadvantage. Thus, unlike most of the earlier research on language for specific purposes (LSP), which is attuned to the native speaker, this study will contribute to this body of literature from another angle: the competent international SLU.

### **1.3 Research objectives**

In this study, the overall research problem addresses developing adequate language proficiency measures that are relevantly tuned to the target purpose and situation of EMI lecturers. The research challenge entails exploring what combination of ingredients point towards a competent international user of academic English for the purpose of lecturing and what language assessment criteria are relevant for capturing this essence.

The dissertation makes research contributions in three areas:

1. The research sets out to determine what qualities of communication should be considered as measures of professional English in international contexts other than the generic, prescriptive rules of the ENL minority. As a by-product, it identifies potential measures relevant to assessing both ELF and LSP.
2. The research pursues language ideologies as a basis for understanding representations of English through self-assessments. How informants talk about their self-assessments is relevant to developing adequate assessment criteria for a given purpose and situation, such as professional English for an international context.
3. A generic L2 assessment tool is examined in relation to the perceived needs of ELF users in order to bring pragmatic value to the research.

To address the overall research question, the investigation utilizes perceptions from three views: self-perceptions, students' perceptions, and EMI-lecturer perceptions of the generalizability of five CEFR scales as a measure of their professional English.

### **1.4 Research questions**

Research on the ideological basis of English language teaching (ELT) practices has already revealed the need to make ELT teaching and testing more relevant to ELF users. In 2000, Jenkins redefined phonological error for communicating in English in international contexts, and in 2007 she advocated against ELF taking the native speaker as a linguistic point of reference (2007: 3). The empirical research in Jenkins (2007), however, investigates the attitudes and identities of ELT professionals primarily in relation to Standard English and NS language ideologies. While this is an important area of research, its implications do not necessarily carry over to non-ELT professionals, who teach their subjects through ELF.

Thus, my first research question aims to examine the language ideologies of non-professional ELT lecturers. In other words, it examines what language ideologies guide the perceptions of EMI lecturers, who teach field-specific subjects in English. Because the professional targets of these two groups differ, we could expect that the perceptions of English language competence for non-ELT professionals may not be primarily based on one language ideology. As non-ELT professionals, the language backgrounds of EMI

lecturers vary widely in terms of their experience in learning and using English. However, what they all have in common is: exposure to teaching methodologies and teaching materials infused with prescriptivism and ENL norms. Since their language learning days at school, however, the world has changed greatly in terms of the users and usages of English, with non-native speakers now being the dominant group. Consequently, this is the audience they face in their classrooms. Lecturing in this context, the research question aims to determine what language ideologies guide the self-perceptions of their professional English.

RQ1: What language ideologies guide EMI lecturers' self-assessments of their professional English?

It can also be assumed that lecturers' positive and negative perceptions of their professional English are related to different language ideologies, as predicted in this hypothesis:

H1: EMI lecturers' language ideologies will guide their positive and negative self-perceptions of their professional English.

The data for RQ1 and H1 were collected through semi-structured interviews with four lecturers, who lecture master's level courses to engineers.

The second research question relates to a language assessment tool. The tool choice is linked to the background of the study. The selected descriptor scales were part of a pilot study for assessing the language proficiency of the EMI lecturers participating in this study (see section 2.1). RQ2 aims to investigate the adequacy of descriptor scales from the CEFR for the target purpose and situation:

RQ2: To what extent are CEFR descriptors a potential tool for assessing the professional English of EMI lecturers?

A related hypothesis is the following:

H2: CEFR descriptor scales do not tap into the skills relevant to professional English for the purpose of lecturing in an international context.

In this part of the study, seven lecturers with engineering backgrounds participated in semi-structured interviews. Using the descriptor scales, the interviewed lecturers self-assessed their English for both professional and non-professional purposes, and then analyzed the relevance of the descriptors in each of the five scales for the purposes of their work.

Not only are lecturers' self-perceptions of interest, but also their students' perceptions of lecture comprehension. Because lecturing is an interactive situation, examining students' perceptions of English is relevant to determining the lecturer's ability to use

language and to communicate content in lectures. This point brings us to the third research question:

RQ3: What are the students' perceptions of their EMI-lecturer's English during lectures, and do their perceptions change during the course?

Through RQ3, the research aims to investigate students' ability to comprehend lecture content as well as their perceptions of their lecturer's English. It also attempts to reveal a relationship between EMI students' English and their lecturer's English. A related hypothesis is the following:

H3: Students' perceptions of their lecture comprehension will change over time as they adjust to their lecturer's spoken English.

This hypothesis is directional and assumes that students will adjust to their lecturer's spoken language over time. For RQ3 and H3, the methodology is based on a quantitative analysis of student perceptions as collected from the same four lecturers in RQ1, with a total of thirty-six matched pair responses.

The treatment of the three research questions and their related hypotheses involves several concepts that need further explanation. Some of these concepts are presented and explained in the theoretical framework in section 1.5.2. Others will be elaborated upon in more detail within the literature review covered in Chapters 2 through 5.

## **1.5 Overall research design**

This section briefly introduces the overall research design, starting with the methodological approach. It then highlights some key aspects of the theoretical framework, and ends with an overview of the dissertation structure.

### **1.5.1 Research methodology**

This dissertation utilizes a mixed-method research design. This means it involves both qualitative and quantitative research, where each part is relatively completed on its own, and then combined to form the basis of one study. Thus, each part of the study is conducted to answer a particular question, and the research results are triangulated to form a comprehensive view (Morse 2003). In this study, the data were gathered concurrently, and the results compared for similar findings.

The quantitative part aims to objectively measure students' perceptions of their lecturer's English while the qualitative part draws on the assumption that perceptions of language competence are socially constructed by individuals and therefore should be investigated interpretively. Thus, the qualitative part of the dissertation primarily utilizes bits from both sociolinguistics and ethnography. The basic assumption is that language is

part of a larger social context and that knowledge about that context is relevant to the linguistic competence of the participants (cf. for instance Rampton 2007). Lecturer language competence is analyzed as part of a social context involving EMI courses and I assume this context affects the participants. My aim is to convey participants' understanding of their professional English for the purpose of lecturing, and to allow their experiences and perceptions to guide the analysis (Blommaert 2007: 682).

In this dissertation, the study of spoken professional English in an international context was approached using mixed-methods. In the qualitative part, one method was to *observe* the primary participants. In practice, this involved my observing selected lectures, and taking field notes. It also included video-recording each participant, which was then used for introspection in a stimulated-recall session. As a complement, the study utilized *semi-structured interviews*, which were open-ended in nature and more oriented towards conversation than a formal interview. To guide the conversation, I employed points of reference from my field observations, the videoed lecture, and selected CEFR scales. The participants in the study included four primary and three secondary EMI lecturers, all of whom have engineering backgrounds. While the primary group participated in the entire study, the secondary group partook in only one aspect of the study, which involved an interview where the discussion was guided by the CEFR scales. In the quantitative part of the study, data were gathered from students attending lectures given by the primary participants. The data include students' perceptions of their lecture comprehension as determined by *initial and post questionnaires*.

### 1.5.2 Theoretical framework

To gain understanding of the data gathered, sociolinguistic theory is applied at an overarching level. I argue that language is *realized* as part of a *heterogeneous* social and cultural *context* rather than as an *idealization* of a *NS* in a *homogeneous* society. Moreover, *standard(ization)* is perceived as more of an *ideological* exercise than a *realization* of any given speaker's output, albeit it has had influential effects on the formal properties of language practices. Taking this as the starting point, I argue for the development of *language assessment criteria* suitable for the context, purpose, and situation of *language users (not learners)*.

Among the skills fundamental to *language competence* for SLU, this research draws on *communicative competence* where *situational appropriateness* does not imply the ability to interact with monolingual NSs (Hymes 1972a). Rather, the aim of speakers is to *accommodate* (Giles 1977) each other as a way to enhance *communicative efficiency* in an *international context*. This communication may include *plurilingual* skills, such as code switching, blending, and mixing. Moreover, these skills are regarded as primary in an international context where plurilingual users dominate. Unlike SLU research, SLA emphasizes SLL competence as modeled on the idealized educated NS (cf. Chomsky 1965) as representative of a monolingual ENL target culture, where deviation from the ideal signals an error. These differences are central to questions related to assessing



language competence, where good testing practices require assessment criteria to be appropriate to the purpose and situation.

Table 1-1 summarizes the key theoretical concepts forming the foundation of the present study. The left side of the table presents the concepts upon which this study draws, and the right side the concepts from which it diverges.

**Table 1-1. Key theoretical concepts that apply to the present study.**

This study aligns with	... and diverges from
<p><b>Language ideologies</b> Perceived as heterogeneous in a given context (<a href="#">e.g. Woolard 1998</a>)</p>	<p><b>Language ideologies</b> Perceived as homogeneous in a given context (<a href="#">e.g. Quirk 1990</a>)</p>
<p><b>Interactionist view</b> Views language as embedded in the context in which it is being used, and where the primary focus is on pragmatics, not grammar. In this approach to the study of language, interlocutors exist in a negotiated arrangement where feedback is possible</p>	<p><b>Idealization view</b> A view that ignores variability in data and that underlies the notion of language competence. It is an assumption in generative linguistics, where the aim is "to account for the language of an 'ideal' speaker in an ideal (i.e. homogenous) speech community, who knows his language perfectly, and is unaffected by memory limitations, distractions, errors, etc., in actually using his language" (<a href="#">Crystal 1985:151</a>)</p>
<p><b>Language variety</b> The term language variety refers to any system of linguistic expression whose use is governed by situational variables, for which several classifications have been proposed, involving such terms as dialect, register, medium, and field (<a href="#">Crystal 1985</a>)</p>	<p><b>Language standard(ization)</b> A prestige variety of language within a speech community that displays no regional variation, thus providing a unified means of communication, that is an 'institutionalized' norm to be used in mass media and in teaching language to foreigners, for example (<a href="#">Crystal 1985</a>)</p>
<p><b>Communicative competence as situational appropriateness</b> Language is related to social codes of speech communities, which entails shared understanding of interactions, norms, and established grounds of authority (<a href="#">e.g. Gumperz 1968; Hymes 1972; Philipsen 1975</a>)</p>	<p><b>Communicative competence as NS ability</b> "The native-speaker's ability to produce and understand sentences which are appropriate to the context in which they occur – what he needs to know in order to communicate effectively in socially distinct settings ... an analogous notion of pragmatic competence" (<a href="#">as defined in Crystal 1985:60</a>)</p>
<p><b>Efficient communication as primary</b> A two-way information sharing process that involves a speaker communicating his/her ideas in a way easily understood by another. To be efficient, speakers may need to adapt their interactions in order to align with or distance themselves from other speakers, in accordance with accommodation theory (see <a href="#">Giles 1977</a>)</p>	<p><b>Linguistic competence as primary</b> The system of linguistic knowledge held by native speakers of a language, which contrasts with the concept of communication. It follows an idealized conception of language that is defined by grammar or a set of language rules that includes phonetics, phonology, morphology, syntax, and semantics</p>
<p><b>Second language user (SLU) (cf <a href="#">Mauranen 2011:160</a>)</b> A 'user' in a contact language situation aiming to communicate effectively to primarily other NNSs, where maintaining their linguistic competence &amp; identity as transportable skills (cf <a href="#">Zimmerman 1998:90</a>) is a natural and appropriate phenomena that may involve code-switching/mixing as a way to promote solidarity and enhance communication</p>	<p><b>Second language acquisition (SLA)</b> A 'learner' acquires language systematically in developmental stages, progressing towards an NS target (regardless of the type of input), although there may be variability while learning and language may show signs of L1 transfer or fossilization (in the event that learning stops) – where deviation from NS norms defines 'error' – a 'deficit' view</p>

As this summary indicates, this dissertation employs a sociolinguistic approach to the study of assessing spoken professional English in an international context where speakers are plurilingual and *variability in language* is present.

## 1.6 Overview of the dissertation

This dissertation contains ten chapters. While Chapter 1 introduces the research topic and the motivation for the study, the next four chapters situate the topic in the relevant literature. Chapter 2 outlines the general context of higher education in a global market, where internationalization is policy-driven and means providing education in ELF (in countries outside the ENL world). The theoretical background is discussed in more detail in Chapters 3-5, each of which focuses on a key aspect: Chapter 3 defines and describes language ideologies relevant to analyzing the data. Chapter 4 discusses language assessment from a performance perspective that includes LSP and ELF. Chapter 5 examines the ideological basis of CEFR. Then, in Chapter 6, the research methods and

design are presented, introducing first the rationale of the study before describing the participants and then explaining the sources of data collection and analysis. Chapters 7, 8, and 9 include analyses of the perceptions of EMI-lecturer language competence from different points of view: as perceived through self-perceptions, as viewed by students, and as seen in relation to five CEFR scales. The final chapter presents the implications and conclusions and then proposes suggestions for future research. A visual representation of the overall structure is presented in Figure 1-1.

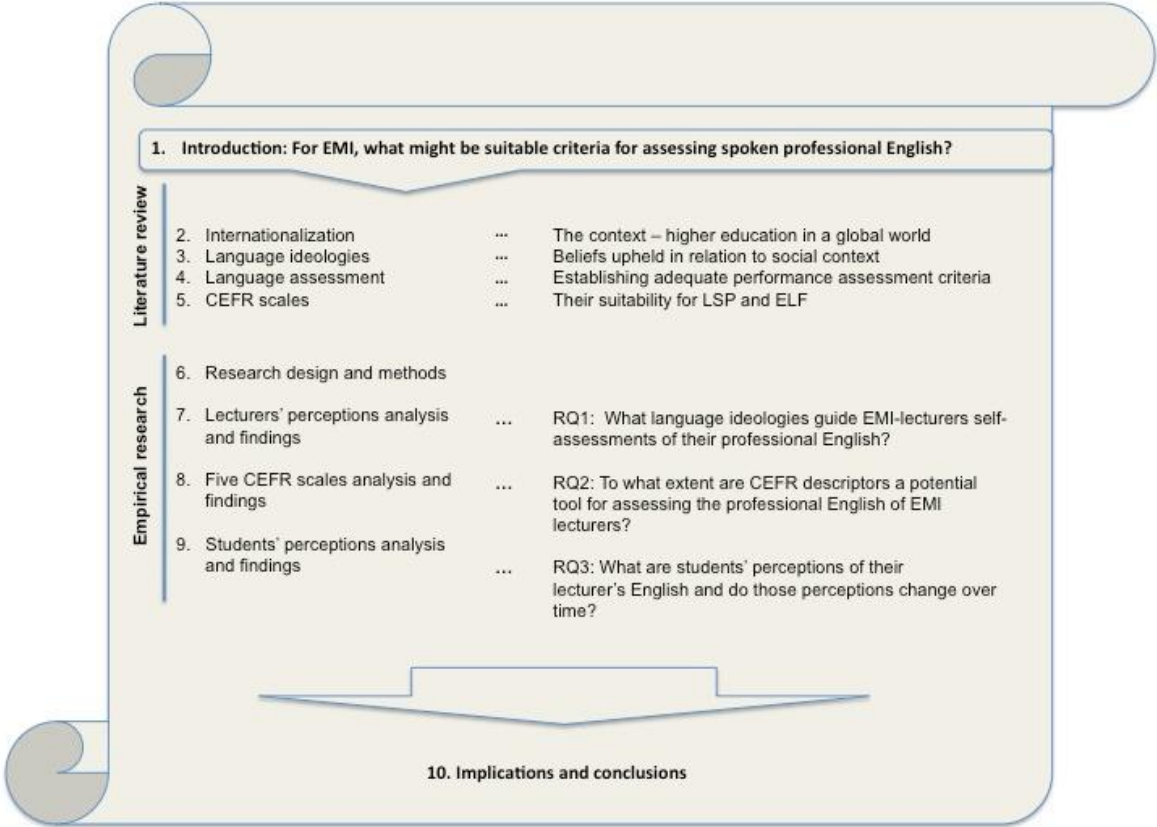


Figure 1-1. A visual representation of the structure of this study.

## **2 Internationalization**

As pointed out in Chapter 1, English is a cornerstone for an international market worldwide, and one that now includes education. The pressure for universities to internationalize has created a huge market for EMI, and with it new job demands for EMI lecturers. To situate the current study in this context, this chapter first presents the background against which the present study transpired. It then reviews literature on EMI to chart the international atmosphere in higher education, after which it situates the study in European language and education policies from the viewpoint of language ideologies. It then defines ELF and concepts related to establishing it, after which it highlights some key research relevant to developing criteria for assessing ELF.

### **2.1 Background to the study**

In this study, the roots of a transformation to teaching through English are firmly planted in university strategies at Aalto University (established 1.1.2010). The strategies include creating an internationally competitive institution, with teaching in English at the master's level in all programs. This plan fits the aim of the Finnish government to extensively reform higher education in Finland (as mandated by the Ministry of Education, OPM 2007).

To give support to faculty teaching in English, a pilot-mentoring program was established (2009 – 2011). During the years 2002-2006, the need for such a program was determined through feedback from students and teaching faculty from major parts of the university: Aalto School of Economics (formerly Helsinki School of Economics) and Aalto School of Science and Engineering, which consists of four schools of technology (formerly Helsinki University of Technology). The feedback from students was critical toward courses taught in English and the need for training had been pointed out by the teaching faculties themselves (see, for example, Suviniitty 2012). Because quality assessment and quality control are regarded as important at Aalto University, the pilot-mentoring program includes in-house language and communication specialists as mentors, who also assess the participants' professional English and pedagogical competence at the end of the program. The assessment, however, was not compulsory. Rather, participants were offered an opportunity to take part in the pilot-certification assessment.

In this pilot-mentoring program, the performance-based certification assessment was developed by an in-house committee of experts, appointed by the pilot-mentoring program management team. This committee consisted of communication specialists from Aalto School of Economics. The aim was to pilot an assessment procedure that could be used to certify the ability of lecturers to teach through English. The aim of the pilot certification was to ensure quality teaching in English at Aalto University.

For evaluating the English of EMI lecturers, the in-house assessment committee developed a test and assessment criteria, which were piloted in spring 2010. Although the test was designed to assess both oral and written skills, the focus was primarily on spoken

skills. The written task was a statement of teaching philosophy to be completed within a two-week time period. The oral tasks consisted of a simulated teaching demonstration with classroom interaction (30 minutes) and an interview (15 minutes). This test design attempted to simulate a lecture with classroom interaction, which includes situational and interactional authenticity. This test task also represents a real-life task. In other words, it was designed to reproduce as closely as possible a target situation that is not just a test of language use. To assess the competence to teach through English, five analytic criteria were used: structure and organization, content, delivery, illustration and support, and proficiency in English. While the test tasks appear to have face validity, I considered the language assessment criteria as questionable. It is this latter aspect that is of primary interest in this dissertation.

For assessing professional English, five CEFR descriptor scales were selected, and the pass level set at C1, except for phonological control, which was set at B2. One reason for choosing the CEFR scales is related to the decision of Language Centres in Finland to adopt CEFR as the measure of language competence for foreign language studies. This adoption was part of adjusting the curricula for language and communication studies to meet the requirements of the Bologna Process (Nikko 2005). For lecturing, the C1 level was considered to be adequate for addressing students. There is no empirical basis for this assumption; instead, it is rooted in pragmatic reasoning. Another reason for adopting CEFR relates to its original purpose: it was created as a reference work with a common core for language teaching, learning, and assessment that would facilitate the comparison of language courses, curricula and proficiency in European countries. Moreover, it was constructed from a heuristic approach, allowing language users (testers, teachers, and learners) to refine it to meet their own needs. In addition, it has become widely accepted as the European standard framework for assessing L2 competence.

At the end of the pilot-mentoring program, participants who passed the performance-based certification assessment were awarded a certificate for teaching in English. This certificate may have a positive impact on their performance evaluation as Aalto plans for teaching in English include linking a certification test to the tenure track system<sup>1</sup>. Moreover, this performance-based pilot-certification assessment represents a first step towards a Finnish university setting targets for a professional language standard for teaching in EMI. As one of the mentors in this program, I was concerned about whether the selected language assessment criteria were suitable for the target purpose and situation, that is, to assess the professional English of EMI lecturers.

## **2.2 Why focus on English-medium instruction?**

At the start of the new millennium, the pressure for universities to internationalize began to radically materialize. Faced with new challenges stemming from globalization, digitalization, and marketability, universities throughout the world began breaking across national borders to enter into the global markets of education and research (Kivistö and

---

<sup>1</sup> *This information is based on an interview with the project leader of the mentoring program.*

Tirronen 2012: 76), which means operating in English – the global lingua franca. In international higher education, offering EMI programs has also been a natural choice since English is the most widely studied foreign language.

In higher education, the key word for internationalization is mobility, which includes students, staff, and institutions (i.e. offshore delivery). Between 2000 and 2004, the estimated number of students registered in programs outside their home country increased more than 40% (Hughes 2008: 3). Hughes also notes a high-end estimation, which suggests 6 million transnational students by 2020 (ibid:1). Currently, universities from four ENL countries (Australia, Canada, United Kingdom, and United States) deliver over 50% of the programs, and universities offering EMI hold a strong position (ibid). Both Europe (see Wächter and Maiworm 2008) and some Asian states (see Kirkpatrick 2011) are striving to become educational ‘hubs’. Although comprehensive statistical data on EMI are lacking, there is common recognition of increasing trends in higher education.

In Europe, a central strategy for connecting higher education lies in the Bologna Declaration. To support and encourage internationalization in higher education, European policy makers harmonized European degree systems in accordance with the Bologna Process (Bologna Process, European Commission 2009). A chief aim of this process was to establish a European Higher Education Area by 2010, making an attractive option for mobility. Since English is the most commonly taught foreign language in the European Union (Eurostat 2012), this harmonization has evolved primarily through programs taught in English. By aligning higher education, European policy makers also aimed to improve the efficiency of tertiary education systems and to advance the idea of accountability (Commission 2007).

This development has promoted international mobility among the academic staff and students in Europe through exchange programs and other types of transnational cooperation. It has also led to an increased number of EMI programs across Europe (Wächter and Maiworm 2008). Moreover, Wächter and Maiworm report Finland ranking as second (just behind The Netherlands) in providing programs taught in English, as measured by the proportion of EMI programs against all programs (ibid: 26). According to Saarinen (2012: 243) the number of study programs in Finland increased from approximately 75 EMI programs in 1996 to 335 international degree programs in 2010 (bachelor’s and master’s level). These programs are primarily in English, with only two in Swedish (a national language) and five in ‘other’ languages (e.g. Finnish and Fenno-Ugric degree programs for foreigners) (ibid). While European universities are becoming increasingly multicultural, the programs that attract international students are primarily conducted in English. Through these programs, students “prepare themselves for an international career” (Mauranen 2007a: 243). In other words, it is not their objective to study English. This situation illustrates the practical use of English as the common language among ELF users having a variety of cultural backgrounds and L1s. Moreover, in this international context, European universities have a commonality: Increased opportunity to participate in exchange programs where the majority of the teaching staff and students have no special relationship with the English language in the sense that they have never lived or stayed for an extended period in an ENL country. With this as the

starting point, the proficiency levels may vary greatly among faculty and students with different L1s who use English as an instrument to communicate.

On institutional mobility, EMI programs are also motivated through a growing trend of satellite campuses. The number of these campuses increased up to 56% from 1999 to 2011 (from 35 to over 200) (Lawton and Katsomitros 2012; Maslen 2009). Leading Europe-origin campuses include the U.K. (13), France (11), and the Netherlands (5) (ibid). According to Maslen (2009), East Asia is the central destination of new campuses where thirty-five branches are scheduled to open in 2012 or 2013. Of these, in Beijing a Danish University Center will open in 2013. With English-taught programs, many of these degree-granting institutions provide another avenue for EMI.

With such patterns of emerging educational mobility, the international mix includes students and lecturers from a wide range of L1 backgrounds. The different mixes indicate that English is and will be widely used as a lingua franca in higher education. This, in turn, means that an increasing number of transnational students will acquire their education through EMI. While there may be a few ENL speakers present, the contact situation is clearly dominated by non-native speakers. In this ELF setting, some speakers may use English in very specific situations, such as managing their studies, and may not need it very much for other aspects of their everyday lives in the local country. To engage themselves more fully in ELF communication, one study even reports exchange students activating a club for ELF speakers (Kalocsai 2009). This club, however, is not extended to local students or to ENL speakers (ibid). This activity illustrates a strong desire to practice ELF communication, and it fosters ELF among plurilingua-cultural speakers. The increased use of ELF in higher education places it in a special position, which has implications for language norms and proficiency measures in international communication.

The general picture that emerges is one where the most important social institutions – the universities and other institutions of higher education – that are the primary sites for the emergence of academic English and social life – now have a strong presence outside the ENL world. In this context, a large number of transnational students are and will be attracted to EMI programs in countries where English is not the native language of the majority of the students or teaching staff. While a few transnational students may be studying in the local languages, the majority conducts their studies in English. As universities offer increasingly courses through English, they in turn need more and more staff that can teach through English. The significance of this should not be underestimated.

Given the strong EMI presence, the communication skills required of EMI lecturers and students should not be interpreted as equaling those of monolingual ENL speakers. The mobility and diversity of EMI participants harbors an environment ripe for innovative plurilingual language use. In this international context, lecturers need to communicate with primarily non-ENL speakers and at an academic level. This means that EMI lecturers need to express conceptually demanding ideas verbally to their students, in a high-stakes educational context, where normative attitudes to standard language may not represent the reality. Being inherently international, academic practices are not closely linked to a national base either (Mauranen 2012). This feature also implies less focus on ENL models (ibid). Similarly, with more scholars using English as an L2, a reduced focus on ENL

norms for ‘good’ English could be expected (ibid). Using professional English in international contexts thus raises questions about the relevance of ENL norms, and gives rise to re-thinking assessment methods. In this non-ENL-speaking matrix, the highest level of proficiency cannot be native-likeness (ibid). What works well in this international context may be very different from appropriateness in an ENL setting. This point is further discussed in sections 2.4.3 through 2.4.5. Since meaningful assessment should reflect the target situation, studying what EMI lecturers say in their self-assessments about their professional English is useful in pointing the way towards developing suitable assessment criteria for this purpose and situation – the central focus of the present study.

Although European institutions of higher education have specified general guidelines for language policy, some are now facing decisions related to establishing adequate language proficiency measures for EMI staff. Against this backdrop, the next section exemplifies education and language policies in Europe as supporting plurilingualism and multilingualism in internationalization strategies for European higher education.

## **2.3 European education: language policies and language ideologies**

Although language policy is not the central focus of this study, I touch on it to illustrate how language policies filter through to assessment practices. I begin by discussing policies of European education at different institutional levels and end with examples of their practical application for assessing the professional English of EMI lecturers. An important component of the discussion is the thread related to language ideologies.

In Europe, both the European Union and the Council of Europe have initiated language and education policies for member states. The policies from the EU encourage language learning and language awareness in the domain of education, as specified in the Maastricht Treaty (Treaty 1992), as well as bi- and multilingualism as specified in the White Paper on Education and Training. According to the White Paper, every European citizen should “develop proficiency in three European languages” (EU 1995: I). These language policies comply with the Lisbon Treaty (2000), known as the Lisbon Strategy. The aim of this strategy was to turn the European economy into one of the most competitive markets in the world through its dynamic knowledge-based system by 2010. This strategy incorporated the ability to sustain economic growth through enhanced social cohesion (ibid). The action plans for this strategy included language learning as one of the thirteen objectives (ibid). These documents recognize language as an important element in building cohesion within society. Building on this foundation, the European Commission published an initiative for promoting multilingualism: *A New Framework Strategy for Multilingualism* (2005a). This framework supplemented an action plan for multilingualism initiated in 2003, *Promoting Language Learning and Linguistic Diversity: An Action Plan 2004-2006* (Commission 2003).

In addition to the aforementioned ideologies of multilingualism and multiculturalism, the policies support plurilingualism. These ideologies are explicit in the aims of the EU for creating composite social identities that regard cultural diversity as an inherent part of

'Europeanness' (Alves and Mendes 2006). They also inform language education policies developed by the Council of Europe. For example, in 2002, the Council published a language policy, which member states could follow, the title of which is *From Linguistic Diversity to Plurilingual Education: Guide for the Development of Language Education Policies in Europe* (Europe 2007, original 2002). The aim of this document was to develop a coherent approach to language education policies in Europe, with a central focus on plurilingualism as a competence and value. In this document, plurilingualism refers to the potential or ability to use several languages, to varying levels of proficiency (ibid).

Another document published by the Council encouraging plurilingualism is the *Common European Framework of Reference for Languages: Learning, Teaching, and Assessment* (2001). While this document supports plurilingualism, it also upholds several other language ideologies (for details, see Chapter 5). In this Framework, plurilingual and pluricultural are defined as "the ability to use languages for the purposes of communication and to take part in intercultural action, where a person, viewed as a social agent, has proficiency, of varying degrees, in several languages and experience of several cultures" (Europe 2001: 168). This phrase denotes a single competence, referring to a speaker who is able to use his/her linguistic repertoire consisting of more than one language. The concept differs from multilingualism, which refers to the presence of more than one language at a particular geographical location. Thus, plurilingualism shifts the focus from the macro context of different languages being present in a social context to those held by an individual speaker. The goal in European language policy is for plurilingualism to be common to all speakers, and for it to form the basis of "intercomprehension" in Europe (Europe 2007: Section I). It is the responsibility of the education systems within each member state to ensure that all citizens are aware of the nature of the ability to use different languages, to emphasize its value, and to develop this skill throughout life (ibid). In this way, European language policies support the formation of a cohesive society, instilled with diverse linguistic and cultural knowledge that will enable individuals to function in diverse situations. This knowledge also plays an important role in internationalization strategies promoted in European education and business.

European language and education policies are increasingly having an impact on decision-making in education. This point is demonstrated in the language policies of European universities, where it is common to promote two or more languages. At the tertiary level, another common expectation is for students and academic staff to be able to perform their duties in English (particularly at the master's level), without any additional preparatory courses than what they obtained in their general foreign language education at school. This point is illustrated in the next section.

### **2.3.1 Higher education in Finland: Language policies**

To meet the aims of the European Higher Education Area, Finland changed its higher education policies and consequently its system, as did many other countries in Europe (Saarinen 2008). To align with this agreement, the Finnish Ministry of Education



published a new internationalization strategy in 2009 (OPM 2009b, The new universities act).

With this international paradigm, the teaching staff is working in a foreign language and the question of how to assess their ability to perform their duties in a foreign language will obviously be dealt with at an institutional level. This point is implicitly confirmed in the strategy documents published by the Ministry of Education, where there is no mention of laws regulating the ‘foreign languages’ to be used in international programs. In the international strategies, the following phrases refer to using ‘foreign languages’ in higher education:

The higher education institutions have increased education in foreign languages leading to a qualification. In proportion to the size of our higher education sector, there is an exceptionally large amount of teaching available in English.

(Ministry of Education 2009:14)

The higher education institutions offer high-quality education focused on their fields of expertise, given in foreign languages.

(ibid: 26)

In Finland, the international strategy for teaching in foreign languages complies with the language policies set by the European Union, which promote linguistic diversity and plurilingualism (see Commission 2005b).

Situating this dissertation in this broad context, the next section outlines the language guidelines at Aalto University. As a new university, the plans for implementing its language guidelines are still in the early stages.

### **2.3.2 Aalto University: Language guidelines**

The language guidelines at Aalto University support the following four principles

- Three working languages: Finnish, Swedish, and English
- Everyone is included in discussion
- Multilingualism is a strength
- Good communication skills are fostered

(University 2010: approved 17 Dec 2010)

These guidelines comply with higher education strategies for internationalization and multiculturalism, set by the Finnish Ministry of Education. They also reflect the degree structure of the university, where undergraduate-level education is conducted in Finnish and Swedish (the national languages) and graduate-level education mainly in English. The language guidelines also promote multilingualism, multiculturalism, communicability, and social inclusion as values, all of which comply with language policies commissioned by the European Union.

The guidelines for teaching in a multilingual environment include the following

- Good language and communication skills are part of the pedagogical requirements for teaching staff
- Those teaching in a language other than their mother tongue are required to demonstrate their proficiency in the language of instruction
- Teachers are provided with pedagogical support and training to enhance their skills in teaching in English in a multicultural academic environment
- International staff members are given the opportunity to study the Finnish and/or Swedish languages

(ibid)

The implementation of these language guidelines is covered by a separate document that contains a five-year plan (2012-2016). This plan outlines and defines the methods through which the measures and resources for a multilingual culture will be implemented. The implementation plan will provide guidelines for a common direction at Aalto University, but will leave room for each School (six in total) as well as individual programs to make their own language decisions (University 2011).

In relation to this dissertation, two guidelines in the five-year plan are particularly relevant. The first one denotes good language and communication skills as part of the pedagogical skills. These skills are also considered an asset in new recruitment. The second guideline of interest requires teaching staff to demonstrate sufficient language proficiency in the language of instruction. Even though this guideline has not yet been activated, Aalto has established a quality assurance system for teaching in English, specified in the guidelines as “teaching through English in a multicultural learning environment” (ibid). This system is being operationalized through an experience-based pilot program, which has already started and is being further developed. This program includes a practical implementation for demonstrating language skills.

In this policy-driven context where quality in teaching is important to the university managerial team, my interest is in how the teaching performance of EMI academic staff will be evaluated. This is a point that obviously concerns university management in many European universities and elsewhere, who are preparing to set standards for teaching in English.

### **2.3.3 From language policy to institutional practices: two examples**

As the previous sections illustrate, language policies are mediated through institutions of social control, such as national governments and institutions, who have the authority to prescribe the representative linguistic values and norms that will then be operationalized. At an institutional level, a project for setting language proficiency guidelines can involve different groups, including student representatives, language experts, pedagogical experts, content teachers, administrative representation, and other educators. For this situation, the concept of language ideologies provides a framework (discussed in Chapter 3) for

considering how institutional authorities enact understandings of language policies in relation to assessing *the ability to function*<sup>2</sup> as an EMI lecturer.

Against this backdrop, the role of language ideologies becomes interesting. The institutional site, serving as a rule provider, may impose a particular language standard, which can be based on different beliefs, such as plurilingualism or standard language, both of which can be central elements in ideological belief systems. Given this situation, a related question is how closely the imposed standards match the reality of communicating in a lingua franca in an international context. To illustrate this point, I will next present two examples of institutional standards set for EMI lecturers at two European universities.

The first example is a large-scale study conducted at Delft University (Netherlands) that assessed the English of scientific staff, which included 1600 participants (Klaassen and Bos 2010). The aim was to determine the average level of the English proficiency of staff and to establish whether IELTS level 6.5 / CEFR level C1 were realistic requirements. The study found that 55% of its staff is at C1 and that they are not achieving a C2 level, which the evaluators consider to be the optimal level for the university to obtain a truly international profile. They further conclude that to prepare students for a global citizenship that they must offer the opportunity to listen to lectures at level C2. Their view to English language proficiency upholds ENL standards (embedded in these standardized tests) as the prevailing criterion for being able to communicate in English in a global world of primarily non-native speakers of English. As the assessment criteria are intended to measure the language competence of SLLs interacting with ENL speakers of a target culture, this view seems to be missing the reality of interacting with plurilingua-cultural speakers in an international context where English is a contact language and few ENL speakers are present, if at all.

The second example of an institution imposing a language standard representative of an illusory community is exemplified in the language certification assessment for teaching staff at the University of Copenhagen, where the target community is very close to the CEFR fallacy of the native speaker being the top-ranking user. For certifying EMI-lecturer language proficiency, the TOEPAS assessment scale employs NS language ideologies. The scale developed includes a five-point global scale, where Kling and Staehr (2011) report the following about the top two levels:

The overall certification result is based on a combined assessment of the lecturer's fluency, pronunciation, vocabulary, grammar and interaction in English for university teaching [where] 5: The lecturer has demonstrated English language proficiency for university teaching equivalent to that of a highly articulate, well-educated native speaker of English [and where] 4: The lecturer has demonstrated excellent English language proficiency for university teaching. The lecturer has been certified to teach English-medium courses. No training is required.

(ibid: 218)

---

<sup>2</sup> This concept originates from the work of Carroll (1961[1972]), and is discussed in section 4.1.1.

As noted in the description, the TOEPAS scale places the native speaker as top ranking, followed by excellent in level 4. This scaling clearly idolizes the native speaker, who is ranked beyond excellent. Such assessment descriptors exemplify native speaker myths of the ‘perfect’ speaker. They also endorse culture-specific features, such as idiomatic expressions, which Pitzel (2009: 300) describes as a “territorial imperative” relevant to ENL membership and Seidlhofer as “unilateral idiomaticity” (2004: 220), both signaling their lack of functionality in ELF communication.

What ENL normative descriptors disregard is the possibility of an ethnocentric communicator who is non-accommodating to SLUs and thus possibly incomprehensible. In other words, speakers of a given L1 are not ‘perfect’ in all aspects of language use - just on the basis of it being their L1 – their communication skills vary, as well as their training for communicating in English with individuals from other cultures. What an ENL-biased scale thus misses is how well an educated native speaker might fare in an international context of plurilingua-cultural speakers using English as a contact language. Moreover, it is hard to imagine how scales like the TOEPAS could realistically capture the proficiency of EMI lecturers who live and work in non-ENL countries. Even if they are native speakers or bi-lingual, they participate regularly in ELF contexts where other ENL speakers are seldom present, and this factor will affect their language – either negatively or positively – depending on the view (see discussion in 4.2.2).

Nor is it clear why the TOEPAS evaluators think a standard based on native-likeness would be appropriate for an international context of primarily SLUs, particularly when previous studies show lecture comprehension difficulties among non-native students attending lectures held by native speakers (e.g. Flowerdew and Miller 1992; Griffiths 1990, 1992; Huang 2004; Kelch 1985). This choice in language assessment criteria exemplifies constructing an assessment based on an illusionary community of ‘perfect’ English speakers who meet the criteria of standard language – regularly used to assess English for ENL target cultures, but which seem distant from the reality of EMI staff and students.

In European institutions of higher education where language committees can define the language proficiency measures for academic staff, the practice of employing standard language and NS language ideologies most likely stems from attitudes similar to those outlined in Jenkins (2007). Jenkins studied the attitudes and beliefs of ELT professionals from several countries, and her findings suggest that ELT professionals adhere to NS and standard language ideologies for the following reasons: 1) the fallacy of ‘standard’ ENL being more widely understood than other varieties regardless of the context of use, and 2) the inherent NS attitude displayed by many ELT professionals as a result of their training, where the goal is to achieve a close affinity to a NS norm. Whether EMI lecturers share the same targets of language proficiency for teaching in English as ELT professionals will be examined as part of the present study.

## 2.4 International communication in English

In this section, the focus shifts from language policies to language use in international communication. My investigation into the professional English of EMI lecturers should be seen in the context of English as a global language as well as a lingua franca. In this section, I will discuss ELF as conceptualized in the literature in order to position my study more fully in an international context.

### 2.4.1 What is English as a lingua franca?

This dissertation adopts a general definition of lingua franca as being a contact language that serves as the means of communication between speakers who have different L1s (Mauranen 2012: 8). Following Mauranen, this definition is situational. Speakers may have different identities in different situations. For example, a speaker may assume an identity related to speaking a lingua franca in one situation, whereas in another the same speaker may assume an identity as a language learner targeting native-like English (Mauranen 2011b). Although a speaker may change identities with the demands of a situation, the default position is that a lingua franca speaker is not an ENL speaker, nor a SLL. In these different situations, speakers may also orient to different norms (see Hynninen 2013). What the literature indicates is that SLLs orient towards the imposed norms of standard language whereas ELF users orient towards natural or spontaneous norms, which arise in groups and communities and which function to increase communicative efficiency (Mauranen 2012: 6). While SLLs strive to achieve native-like proficiency targets as deemed appropriate by standard language, SLUs have no target culture (ibid). In ELF, language is a means of communication where the primary aim is to be understood (Jenkins 2000; Seidlhofer 2001).

The term, English as a lingua franca, has become commonly accepted for what was earlier referred to, and sometimes still is, as English as an international language, English as a world language, or English as a global language (Seidlhofer 2004: 210). Unlike the other terms, English as a lingua franca, contains no descriptor connecting English to contexts extending beyond national borders. Moreover, the term captures the historical meaning of a lingua franca<sup>3</sup>, thereby freeing this type of English from the idea of NS ownership. It thus shifts the focus to NNSs as the primary communicators. In this way, the

---

<sup>3</sup> *The term 'lingua franca' originates from its meaning in Arabic (Faranji) and Greek (Phrankoi). Its usage comes from Arab-speaking traders in the 1200s, who needed to communicate with Western Europeans and others with whom they shared no common language. The Arabs referred to these non-Arabic-speaking traders as 'Franks'. They also called the language that developed as their trade language 'lingua franca', meaning the language of the Franks. This language resembled a pidgin in that it was not a native language for anyone and it was formed for a specific communicative purpose (Adler 1977). Although other notions of its origins exist, this explanation approximately describes its genesis. Since its birth, the use of the term has slightly changed: it is now used to refer to native languages that serve as 'instrument' languages for communicating in situations where speakers have no other common language.*

term differs from some previous terms, such as English as an international language, where communication takes place in an international setting among interlocutors (native and non-native) who aim to follow the norms of Standard English.

Moreover, it is the large-scale use of ELF that makes it "the first truly global lingua franca" (Mauranen 2012: 17). Accompanying the large-scale use of English is a change in the uses and users of English, as the number of non-native speakers exceeds the number of native speakers (Crystal 2003; Graddol 2006; Jenkins 2000). Even though by definition a lingua franca has no native speakers, it does not mean that it is restricted to a pidgin language. On the contrary, ELF includes a range of speakers, some of which use sophisticated language in numerous, often influential, domains. However, even those among the sophisticated users of ELF have considerable linguistic variation due to the diverse linguistic backgrounds. Diversity means that there is quite a spectrum in mutual comprehensibility. It also means that speakers may need to apply proactive communicative skills for managing ELF interaction, such as clarification and repair strategies that assist in promoting communication (Mauranen 2006a: 147). While dealing with diversity involves primarily SLUs, it does not exclude native speakers – as they may also be present in an international context. Thus, diversity refers to the whole spectrum of users of English.

This spectrum includes speakers of first, second, and foreign languages, a division widely accepted in the literature and mirrored in Kachru's concentric circles model (1985). This model places countries in either the inner circle for which English is the primary or native language (e.g. U.K., U.S.), the outer circle for which English is a second language (e.g. Hong Kong, India), or the expanding circle for which English is used as a foreign language but has no special administrative status (e.g. Finland, Japan). However, this model is not without problems. For example, the distinction between the outer and expanding circles is not exactly clear-cut as these two categories share some characteristics (Kachru 1985: 13-14). Moreover, the circles have become more blurred with the massive spread of English around the world (Graddol 2006: 110), a phenomenon that the model does not reflect. Missing from Kachru's model is the more recent way of conceptualizing English in global use, which is to speak of ELF, a category that could be seen as cutting across all three circles. This, however, is not surprising, given that his model is thirty years old.

#### **2.4.2 NS ownership controversy**

While the rapid spread of ELF has created a new research area for linguists, the views to ELF are not without conflict. The differing views evolve around two main themes: the NS ownership of English and the teaching and assessment of L2 English. Some linguists argue that ELF is a legitimate variety in its own right, while others prefer to ignore it or to attribute any relevance to it (see for instance Jenkins 2007 for an extensive discussion). Central to the NS ownership theme is the question of who has the authority to make decisions about standards for English. Some key publications contributing to the conflicting views can be traced to Quirk (1985) and Kachru (1985), as well as to

Widdowson (1994). While Quirk and Kachru argued over standards to be used by different users of English, Widdowson asserted the need to recognize that no nation has control over the users and usage of an international language (1994: 389).

At the start of the new millennium, two seminal works appeared that added impetus to the controversy. The first was Jenkins (2000), an empirical study of ELF pronunciation where she argued that native-like pronunciation is not optimal in ELF interaction (further discussed in 2.4.3). She also raised a relevant question about standards setting for English:

No one denies the 'rights' of so-called 'native speakers' to establish their own standards for use in interaction with other 'native speakers' [ENL], and even with 'non-native speakers' (EFL). However, the important question is: who should make such decisions for communication wholly between 'non-native speakers', i.e. for English as an International Language [i.e. ELF]?

(Jenkins 2000: 7)

This question is highly relevant to international contexts, where the needs of SLUs differ from SLLs.

The second seminal work was Seidlhofer (2001), who convincingly argued that ELF was "the most extensive contemporary use of English worldwide" (ibid: 133). She also pointed out the lack of description available for this linguistic reality, which "preclude[d] us from conceiving of speakers of lingua franca English as language users in their own right" (ibid) and this meant that the native speaker continued to be held as the only authoritative target for non-native speakers. The position of ELF scholars is that English does not belong to native speakers. Nor are ELF users language learners. This, however, does not mean that language learning does not take place in ELF interaction (see Firth 2009) or that ELF users do not at times see themselves as learners (see Mauranen 2012). The central premise in ELF research is that SLUs communicate in English to accomplish tasks and as such are not learners of the language. This position differs from SLA research, a point further discussed in 2.4.3 and 2.4.4.

### **2.4.3 Idealized NS or competent SLU?**

Much of the controversy over the norms for English include pedagogical concerns related to a NS model. Prominent scholars question the use of a NS model for all learners and users of English (e.g. Jenkins 2000, 2006b, 2007; Mauranen 2003; Seidlhofer 2001, 2004). Moreover, for NNSs to reach the level of native-likeness is exceptional (see for example Marx 2002). While it seems reasonable to distinguish between ELF and ENL, the linguistic accuracy of SLLs is central to a NS model. Moreover, it is rooted in the concept of an abstract, monolingual NS in a target culture rather than what is relevant to successful communication among plurilingual speakers using ELF in a non-target culture. The inappropriacy of the NS view for SLUs is best captured in the words of Cook, who writes "L2 users are not the same as monolinguals" (2001: 177). He further argues "[w]hat we need [instead] is a model that recognizes the distinctive nature of knowing two or more

languages and does not measure L2 knowledge by a monolingual standard” (ibid: 194). This notion suggests a plurilingua-cultural competent user as the ideal against which to model L2 proficiency.

Shifting from a NS model of proficiency to a competent international SLU would mean that ELF users would not need to demonstrate NS standard language for lexicogrammatical accuracy, native-like pronunciation, and other NS qualities of linguistic behavior. Instead, the target would be a competent SLU of English<sup>4</sup> who can operate between languages and cultures (see Cook 2001). Such a shift also implies an alternative paradigm of language proficiency as the fundamental differences point to contrasting qualities as important features defining successful communication in different macro-level contexts: culture-specific versus lingua franca. Central to these two paradigms are the concepts of SLA and SLU, the topic of the next section.

#### **2.4.4 SLU vs SLA**

Here, it would be important to note that the research objectives of SLU differ from those of SLA, and this difference is complementary. In other words, SLU is an additional research area, not a replacement for SLA. Briefly, I will highlight a few of the key differences. The first is the context of language use. SLU research investigates the natural language of ELF users who use English in diverse, international contexts. This research is largely based on spoken corpora (Mauranen 2011b), whereas SLA research studies the language of SLLs, largely under experimental conditions, in monolingual ENL contexts. Another difference between the two paradigms is the SLA target culture versus SLU non-target culture dimension. On this point, Jenkins (2006a: 152) criticizes the SLA paradigm as having a monolingual NS bias, which renders it inapplicable to ELF.

A third major difference in these two paradigms is the treatment of SLL or SLU deviations from ENL, viewed as an error in the former and as a possible ELF variant in the latter (see, for example, Mauranen 2012; Ranta 2009). In SLA, it is common to study SLL from a cognitive perspective and through interlanguage theories, approaches irrelevant to ELF (Jenkins 2006a: 150). This approach takes a negative view to language variation, deeming any language seen as non-standard as ‘deficient’. This view can also be seen from different perspectives: how you see another speaker’s language if you set out to examine it, and how you see your own language as part of your self-image. From the perspective of an examiner, an evaluation requires applying a set of linguistic criteria that deems non-native features as deviant, whereas the perspective of self-image can involve what is known as “linguistic insecurity” (cf. Labov 1972). The latter concept describes a speaker holding negative self-perceptions of his/her own speech variety based on the idea of lack of correctness – often the view of SLLs who see themselves as having failed or

---

<sup>4</sup> Different names have been proposed for this alternative concept, including the following: Jenkins (2000: 9-10) suggested *bilingual English speaker (BES)* and *(NBES) non-bilingual English speaker (who is bilingual but not in English)* contrasted against *monolingual English speaker (MES)*; and Smith (2010:51) introduced *multilingual English speaker (MuES)*.



then lacking confidence in the skills acquired. In contrast, a positive view sees variation as natural (e.g. plurilingualism) and accommodation as integrational behavior. This view is present in Giles and Coupland's (1991) Communication Accommodation Theory<sup>5</sup>. This theory accounts for two main accommodation processes, convergence (e.g. accommodation) and divergence (e.g. over- and under-accommodating), as motivations behind individuals adjusting their speech styles in an effort to reduce social distance between themselves and other interlocutors. In ELF, accommodation skills have been found to be essential, a point further discussed in the next section.

#### 2.4.5 Some key findings in ELF research

By the turn of the millennium, English was increasingly being used and in lingua franca contexts outside ENL countries. This increased usage called for a focus on ELF, and research interest into it began to flourish. At about that time, two large corpora of spoken ELF were launched, thereby providing a great resource for empirical research. In 2001, Seidlhofer launched the Vienna-Oxford International Corpus of English (VOICE), under her directorship. The same year, Mauranen launched the English as a Lingua Franca in Academic Settings (ELFA) corpus, under her directorship (see Mauranen 2003). The former was undertaken to provide a basis for describing the general use of spoken ELF in various settings (i.e. leisure, academic, and professional), and the latter specifically for describing usage in academic settings.

The central aim of ELF research is to describe ELF, the results of which will provide understanding of its use in international communication. The point of departure for determining standards against which to assess ELF will thus be based on empirical research of discourse communities that use ELF. On this point, ELF diverges from L2 performance assessment where the measure of linguistic correctness has been, almost exclusively, the idealized NS. Shifting away from a NS model means taking another view to 'error'. It also means re-conceptualizing a number of other concepts related to a SLU paradigm (discussed in 4.2). In this section, I will discuss some key findings relevant to developing criteria for assessing ELF.

From the point of view of phonology, Jenkins' (2000) work is groundbreaking. Based on her research, she proposed a Lingua Franca Core, arguing that intelligibility is more important for ELF users than imitating an ENL variety. Her research shows variation in ELF communication to be unproblematic as long as speakers can produce the core sounds (for a list, see Jenkins 2000: 159). Non-core sounds found to be irrelevant to intelligibility include the theta sounds, the dark lateral allophone, weak forms and certain prosodic features (see the list, *ibid*). She also reminds the reader that listener factors are just as important as speaker factors (*ibid*). In other words, communication is an interaction

---

<sup>5</sup> *Communication Accommodation Theory stems from Speech Accommodation Theory, both of which originate in the work of Giles, starting with his study on accent convergence in 1973. Giles, Howard 1973. Accent mobility: A model and some data. Anthropological Linguistics, 33, 27-42. For a summary of the expansion of this theory, see Gallois, Ogay and Giles (2005) in Gudykunst pp. 121-148.*

involving both interlocutors, where the success depends on their mutual intelligibility as well as their ability to adapt and accommodate each other. Jenkins' work, however, has been viewed as controversial, and thus is not without its critics. Some ELT professionals have expressed concerns about teaching pronunciation without a NS model. For example, questions have been raised about what impressions a heavy accent might make or the effects it might create and whether an accent that belongs to no one should be taught (see, for instance, Harris 2002; Sobkowiak 2005; Szypra-Kozłowska 2003).

Investigation into ELF accommodation reveals features central to communicative efficiency. Key to this is strategies of cooperativeness and co-construction. One striking feature in ELF cooperativeness is the use of proactive strategies to prevent and manage misunderstanding. Mauranen's (2006a) study highlights this and shows how it emerges unprompted by prior indication of misunderstanding. The proactive strategies include clarifications, confirmation checks, self-repair, and interactive repair (ibid: 135). An example of a self-repair common in her data is a false start. Self-repairs are "strongly proactive" and as such not explicit to intelligibility, but do exemplify heightened language awareness. Interactive repair, on the other hand, is dialogic and can occur preemptively or retroactively. It can thus be used as a preventive measure or as a response after a problem has surfaced (ibid: 137). On accommodation, Mauranen also emphasizes the negotiation of approximations described as "speakers changing their language to approximate what they believe to be the patterns of another language or dialect" (ibid: 126). Her research shows how these strategies enhance explicitness in ELF communication and thereby minimize misunderstandings (Mauranen 2006a, 2012). Other studies also show accommodation skills as essential in ELF communication and as serving a number of functions: enhancing intelligibility (Jenkins 2000), increasing communicative efficiency (Cogo and Dewey 2006), establishing a community of practice (Ehrenreich 2009), identifying with an ELF community (Cogo 2009), and comprehending expression (Rogerson-Revell 2010; Smit 2008; Sweeney and Hua 2010).

In a study on collaborative discourse, House (2003: 569) found features such as co-construction to be very frequent and concluded that it could be the most important aspect. The literature indicates cooperativeness as a common manifestation in ELF data, and as consisting of strategies for enhancing ELF communication, such as the "let it pass" principle (Firth 1990) or the "make it normal" principle (Firth 1996). While communication among monolingual ENL speakers also involves cooperativeness, this can include culture-specific features such as idiomatic language and colloquialisms – likely to result in miscommunication in an international context. Thus, context factors into successful strategies as well as heightened self-awareness of linguistic variation and openness to approximations.

Studies on the context in which ELF communication takes place emphasize diversity and local relevance as important factors that affect the question of norms. For example, Knapp and Meierkord (2002: 10) describe the macro-context of lingua franca communication to be diverse and heterogeneous. Another study examined the 'habitat' factor (cf. Bourdieu 1991) and showed the place of interaction to be a decisive factor for cultural identity in ELF interactions (Pözl and Seidlhofer 2006: 155). The main finding showed local pragmatic norms to apply in ELF interaction when speakers are primarily

from one culture and in their homeland. They argue that the habitat factor has a socio-psychological effect on the speakers (ibid: 172). Similarly, other studies show that ELF users do not necessarily conform to ENL norms (Hynninen 2013, Mauranen 2012). Some studies report that speakers generally cannot depend on their L1 norms due to multilingual heterogeneity (see for example House 1999 and Meierkord 2000). What these different findings suggest is that the context of use is relevant to shaping sociolinguistic behavior in ELF communication.

On lexicogrammatical features, a recent study that examines spoken ELF in academic contexts found it to resemble closely spoken academic ENL (Mauranen 2012). The comparison was based on recorded spoken data from ELFA and MICASE (i.e. Michigan Corpus of Academic Spoken English). While this finding may sound surprising, the study also points out that there are also clear differences between ELF and ENL. Obvious differences manifest as minor variations in form, which present themselves "in approximations, rephrasings, and non-standard syntax, together with apparent dysfluencies in execution, such as frequent repeats, hesitations, and pauses" (ibid: 248). Another difference is the phonological aspect, which was not included in her study. As Mauranen points out, while there are similarities, these features easily distinguish between the two (ibid). Earlier ELF studies have shown that typical errors do not necessarily hinder communication, such as omitting the third person -s, using articles and prepositions inconsistently, mixing relative pronouns (i.e. who and which), simplifying tag markers, overusing generic verbs (e.g. do, make, have), and increased explicitness (e.g. green color) (Seidlhofer 2004: 220). On the other hand, expressions that are ENL tied may create problems, such as metaphors, idioms, and other fixed expressions (ibid). While these features differ from ENL and thus would carry negative consequences for SLL assessment, their role in ELF communication may serve a useful purpose and thus should not be dismissed as such.

## 2.5 Summary

The chapter first described the background to the study, and then focused on the importance of EMI in a global market of education in which tertiary institutions are currently engaged. It outlined European language and education policies and then argued that in practice they filtered into assessment practices based on language ideologies that are not well suited to the context of use. As European universities are moving towards implementing measures for assessing the quality of teaching in EMI programs, a key question is what language assessment criteria might adequately reflect the realities of international communication in English for this context and purpose.

ELF was also discussed as the language of international communication. The review drew attention to two opposing mainstream paradigms (SLA vs SLU) and contentious issues related to NS ownership. It also highlighted some key findings relevant to re-conceptualizing competences related to ELF communication, including aspects of phonological and lexicogrammatical usage as well as accommodation. This literature is

relevant to the development of assessment criteria for professional English in international contexts.

### 3 Language ideologies and their social position

The central focus of the first research question is language ideologies as guiding principles to assessments of EMI-lecturer language proficiency. The theoretical framework presented in this chapter will serve as the basis for analyzing the data in Chapters 7 and 9. The framework consists of three parts. The first part introduces the notion of language ideologies. In the second part, I outline my working definition of language ideologies. The third part describes four language ideologies relevant to the theoretical framework and the final section concludes the chapter.

#### 3.1 The notion of language ideologies

The study of language ideologies evolved from linguistic paradigms related to language shift, language planning, and linguistic anthropology, and provided a means for interpreting notions about language and for analyzing the collective sociolinguistic behavior of a given group (Blackledge 2000: 26). As an analytical resource, language ideologies opened a new chapter for the study of language. With it came an emphasis on the perceptions that individuals held about language and discourse and about how these connect to social phenomena, a notion that had previously been marginalized in the study of language. As Paul Kroskrity contends “this surgical removal of language from context produced an amputated ‘language’ that was the preferred object of the language sciences for most of the twentieth century“ (2000: 5).

Although the initial research focus in language ideologies was related to equating a given language with a given speech community and establishing links between national or regional groups and their linguistic practices, the research focus shifted in the latter half of the 20th century to include nuances, such as social positioning, bias, and variability, that discern language use and beliefs connected to power relations and political agendas in societies (e.g. Blackledge 2000; Blommaert 1999; Blommaert and Verschueren 1998; Gal 1998; Gal and Woolard 1995; Kroskrity 1998). With the shift in research focus, a diversity of work manifested with a range of definitions of the term “language ideologies”. These definitions each highlight different foci illustrating the application of the proposed concept. In this dissertation, I steer away from notions of language such as that proposed by de Saussure<sup>6</sup> (Meisel and Saussy 1959), which implies linguistic homogeneity within a

---

<sup>6</sup> *In Course in General Linguistics, de Saussure acknowledged in ‘parole’ that individual language systems are partly idiosyncratic. However, the existence of ‘langue’ was superimposed and understood as “a grammatical system that has a potential existence in each brain, or, more specifically, in the brains of a group of individuals” (Meisel and Saussy 1959:14). From this follows the assumption that a linguistic community is bound by a common language system, which by definition means that a language community consists of a group of individuals united by a shared language. Here the connection between language and community is a contingent fact that stems from conventions or agreement. This synchronic system assumes members of a given linguistic community to be unified by factual linguistic homogeneity.*

given community of speakers. I also depart from definitions such as that put forward by Alan Rumsey<sup>7</sup> (1990), which imply that language ideologies within a given social group can be viewed as homogeneous. Instead, I adopt Woolard's formulation of language ideologies as, "representations, whether explicit or implicit, that construe the intersection of language and human beings in a social world" (Woolard 1998: 3), as this definition calls attention to the heterogeneity of ideological positions in a given context and implies that these representations can approximate conscious awareness.

While many definitions of language ideology have been proposed, here I will quote two that emphatically restore the importance of the *contextual factors*. An early definition by Michael Silverstein describes language ideologies as "sets of beliefs about language articulated by users as a rationalization or justification of perceived language structure and use" (Silverstein 1979: 193). In this definition, Silverstein mentions the *explicit* expression of language ideology as a voiced rationalization pertaining to the social context<sup>8</sup>. Another definition that brings out the contextual features more explicitly is one proposed by Judith Irvine. She defines language ideologies as "the cultural system of ideas about social and linguistic relationships, together with their loading of moral and political interests" (Irvine 1989: 255). While the foci of these definitions are different, both seem to contain an element of local cultural knowledge as pertinent although they do not in themselves spell this out.

As indicated in the above definitions, the research foci may vary when studying language ideologies. Thus, unlike Silverstein's early definition that mentions *explicit expression*, Jaworski and Coupland (2004: 36-37) focus on the *implicit assumptions*, where language is used against a backdrop of ideological assumptions that are evaluative and prescriptive in nature. In their definition, implicit assumptions make reference to correctness, appropriateness, permissibility, and other similar evaluative words (ibid). They argue that these assumptions are "part of the specific socio-cultural frames, with particular histories, tied in to particular power struggles and patterns of dominance" (ibid: 37). This statement suggests that the perceptions are connected to contentious issues related to power and dominance as socioculturally situated in their particular histories, signaling the prominent position of language ideologies on language use.

The influential position of language ideologies in relation to language use is one focal dimension in the literature, highlighting *power relations and patterns of dominance*. For example, Blommaert (2006: 510) argues that language ideology is connected to the exercise of power and dominance whereas Woolard (1998: 6-7) mentions that language ideology is used to obtain (and maintain) power. In addition, Gal (1998) points out that the power in language ideologies resides in the ability to ascribe value to the practices of language forms used by one group over those used by another. On power relations, Seargeant (2009: 27) argues that language ideologies "reproduce hierarchies within

---

<sup>7</sup> "Shared bodies of commonsense notions of the nature of language in the world" (Rumsey 1990: 346).

<sup>8</sup> While this definition is one of the earliest, a predecessor to research on what is now known as language ideologies was Hymes' work on the ethnography of communication (1974). It was his work that suggested that an ethnographic account include a speech community's own theory of language use (ibid: 31).

societies“. Examples of the semiotic processes (cf. Irvine and Gal 2000) that can lead to indexing and hierarchical relationships are provided in 3.2.

On dominance, a term important to the present discussion is *hegemony* (Gramsci 1971), a concept developed from the notion of dominance. It has come to be associated with dominant discourse, which may be invisible and taken-for-granted, that creates symbolic power controlled by the state. Gramsci, however, does not believe that the control of a dominant position can be maintained over time without the consent of a formal political entity. He views hegemony as being related to both domination and integration. In other words, hegemony is about the process of a dominant group exercising power over an entire society, which requires making alliances with subordinated groups (Fairclough 1995). In Simon (1982: 21), the term is succinctly summarized as ‘the organization of consent‘. This consent pertains to certain ways of viewing the world and making sense of it, which may necessitate adopting an ideology that appears as a natural, common sense choice. Some will opt for the common sense view while others will not, thus creating struggles and at different levels in society, including local (for example, family, workplace), national (for example, educational policy), and international (for example, multinational corporations). While hegemony is a discernible process, it is not stable or colossal. Rather, it shifts, being produced and reproduced, and is characterized by contradiction, opposing identities in populations, as well as counter hegemonies (Blommaert 1999, Gal 1998). In an increasingly globalized world, however, hegemonic discourse at the level of nation-state may no longer be the main actor – if it is an actor.

The instrumental use of language ideologies in shaping language use also extends to definitions that focus on *linguistic differentiation*. For example, Gal and Irvine (1995) define this as

[T]he ideas with which participants frame their understanding of linguistic varieties and the differences among them, and map those understandings onto people, events, and activities that are significant to them ... we call these conceptual organizations ideologies because they are suffused with the political and moral issues pervading the particular sociolinguistic field, and because they are subject to the interests of their bearers‘ social position.

(1995: 970, emphasis in original)

This definition conceptualizes language ideologies as a structure from which to understand linguistic differences, also known as variation. A clear example of such differentiation is present in the ideological motivations behind Standard English in the United States and Britain. For instance, in the U.S. Standard English evolves around lexical and morphosyntactic structures related to the concept of a mainstream, non-ethnic, middle-class speaker, whereas in the U.K. Standard English is associated more closely with the phonological features of a highly educated upper class speaker. In both cases, there is an attempt to erase (cf. Irvine and Gal 2000) the ethnic differentiation, but through different motivations. These examples illustrate the varying ideological bases upon which a standard language can be formed, as constructed in the interest of a particular group (Kroskity 2000). In addition, previous research reveals that these varying ideological

positions may result in a given form being accepted as standard language in one variety, but not in another.

While the definitions presented thus far all have different foci, they also share a commonality: the tradition of language ideological research having a close affinity to monolingual language ideologies where one ideology is dominant. In more recent years, however, the literature shifts in focus to language ideologies and discourse practices as being both *context-dependent* and *cluster-oriented*. For example, Kroskrity (2004: 498) describes language ideologies as “beliefs, or feelings, about languages as used in their social world“. This view suggests that language ideologies are context dependent. Moreover, grounding them in social experience means that the experiences are not uniformly distributed or homogeneous. This view to language ideologies is in opposition to views that conceptualize language ideologies as single shared perceptions of language in a homogeneous community. A theoretical focus on *multiplicity* presumes the existence of different language ideologies within a speech community. A benefit to adopting multiplicity for the study of language ideology is that it permits a focus on potential conflicts, their formulation, and the implementation of a dominant language ideology, for example (Kroskrity 2000: 12-13). Whether a language ideology is contentious or neutral, the point is that a focus on the dimension of *interest* can stimulate perspectives on sociocultural analysis through rethinking cultural explanations relevant to the association of local practices and interests in specific discourse (Kroskrity 2000: 8-12). Thus, while members of a group share language ideologies with similar interests, the same ideologies are not shared by an entire speech community. This multiplicity dimension also means that an individual can be guided by more than one language ideology, and even conflicting ones (see, for example, Briggs 1996; Gal 1993; Urciuoli 1991, 1996).

In short, the literature shows a broad range of notions about language ideologies as different researchers emphasize different *foci*. What is clear from the definitions and notions presented here is that language ideologies inform the beliefs of language use and may have an impact on users' social actions. Because language ideologies are connected to the beliefs and experiences of both individuals and groups as related to their social contexts, this suggests that multiple language ideologies are likely to be present within an individual as well as within different speech communities. Following this view leads to the conclusion that it is beneficial to view language ideologies as multiple since this view acknowledges the different interests connected to language ideologies that inform the linguistic and social behavior of both individuals and groups. This view also explains the presence of shared and/or conflicting language ideologies. As to prevailing language ideologies, the power relations and patterns of dominance in a given social context will guide this. Because the role of language ideologies is so closely tied to social context, this aspect also merits attention and thus is the focus of the next section.



### 3.2 Language ideologies interfacing with social action and structure

According to the above notions, language ideologies guide what individuals say about language use as well as how they use language, and their ideologies are subject to power relations and dominance that raise questions about social justice. Thus, they are about moral or political interests, and they are socially situated through ties to questions about identity and power. In this section, I will further explore language ideologies while illustrating more closely their ties to social action and structure. It is because of these close ties that I included them in the present study. Of specific interest is how language ideologies inform individuals in their social practices in given contexts.

Recent empirical research suggests that language ideologies structure social behavior. This point is evident in Woolard's introduction to *Language Ideologies* (Schieffelin et al. 1998), where she describes the target of the book to be not just an analysis of the social origin of language ideologies but also an investigation of their effectiveness in "the way they transform the material reality they comment on" (Woolard 1998: 11). Even though the Silverstein (1979) definition in section 3.1 indicates that language ideologies may be overtly articulated by group members, researchers are also aware that language ideologies may be interpreted from the actual usage. This point is also present in Giddens' (1984) structuration theory. According to his theory, social systems are created and (re)produced based on an analysis of structure and agency, where both aspects are treated as balanced. More specifically, his theory examines social practices at the intersection of structure and agency, where the two aspects are inseparable. The central tenets of structuration theory maintain that an analysis of society requires both subjective (agent-focused) and objective (structure-focused) investigation. His theory further postulates that an analysis cannot be based exclusively on micro- or macro-focused interpretations. In other words, the connection between social agents and linkage to structures (micro and macro) takes into consideration differing degrees of members' awareness of their own rule-governed activities, including discursive and practical consciousness. For instance, the hierarchical relationship between a student and a teacher is preserved even when they happen to coincidentally meet in a grocery store. This example illustrates a relationship developed in a social structure (i.e. an educational institution) that exists independent of the context in which it was created.

As mediators between social structures and forms of spoken language, language ideologies bind sociocultural experience with their linguistic discourse "by constituting those linguistic and discursive forms as indexically tied to features of their sociocultural experience" (Kroskrity 2000: 21). This mediating role of language ideologies is evident in Irvine and Gal (2000), who posit three productive semiotic processes as underlying the reasoning: iconization, recursiveness, and erasure. Iconization is a process whereby a linguistic feature is pegged to index a social group and thereby becomes iconic. In other words, it is a feature considered to denote the inherent nature of a particular group. As an example, I will use the concept of the native speaker. As a concept, the NS supports the belief in a close relationship between being a NS of the national language and being accepted as a citizen in its linguistic community. This ideological representation can be

construed through iconization. The process involves indexing groups according to whether they speak the standard language of the nation-state, an action that iconizes the standard language as an ideal representation of the whole society. Closely related to this ideology is erasure, which occurs when sociolinguistic phenomena that are inconsistent with an ideological representation are overlooked or glossed over (ibid: 65). For instance, the notion that language is homogeneous in a monolingual speech community permits “a rigid and clear distinction between being a native speaker and not being“ (Pennycook 1994: 176). Through erasure of activity that is different, this language ideology either transforms all individuals to match the target or renders them invisible. This process erases diversity in the linguistic practices of individuals, resulting in the notion of a homogeneous speech community. As an example of the third process, recursivity, by adopting a standard language that automatically designates a high level of language competence in the L1 (in all domains), NSs create a sociocultural framework for interpreting linguistic difference at the level of standard versus non-standard speakers, which is then projected onto another level that leads to creating a difference between NSs and NNSs (see Bourdieu 1991).

Fairclough (1989 [2001]) also discusses how language is used to construct ideologies and ideological positions. A question raised in his research relates to the extent to which *language as code constrains our thinking*. His view emphasizes ‘common sense’ assumptions where authority and hierarchy are treated as natural perceptions of a situation. He points out that “the exercise of power, in modern society, is increasingly achieved through ideology, and more particularly through the ideological workings of language.” (1989 [2001]: 2). This implies that language ideologies constrain our thinking. An example could be how the conventions of teaching an English lesson between teachers and students embody ‘common sense’ assumptions, that is, the teacher knows about grammar and the student does not; the teacher is in a position to determine how a grammar problem is dealt with and the student is not; it is ‘natural’ that the teacher should make the decision and control the treatment of errors, and the student should comply and cooperate and so on. Not only does the authority of the teacher play a role in shaping language ideologies, but also the prescriptive materials distributed in class. Consequently, learners are fed certain ideologies that shape their perceptions of their language, themselves, and the world around them.

In conclusion, language ideologies interplay with linguistic behavior and social structure, and this function involves complex processes. To gain understanding of such complexities requires investigating the ideological basis of language as perceived and used by its social agents in the social structures where the social interaction occurs. For this reason, an investigation into the language ideologies of EMI lecturers at Aalto University has been included in the present study.

### **3.3 Language ideologies as gatekeepers**

As discussed in the previous sections, language ideologies intersect with social action and structure, where power and dominance issues are related to moral and political interests. In recent empirical work on language ideologies, Jaworski et al. (2004) capture this essence

in the following statement “it is in the interplay between usage and social evaluation that much of the social ‘work‘ of language – including pressures towards social integration and division, and the policing of social boundaries generally – is done“ (ibid: 3). In this section, I will examine the evaluative function of language ideologies as a social control of language use, known as gatekeeping.

In the literature, the term gatekeeping has been used to refer to the social values and political agendas of the decision-making bodies that control the standards of language proficiency and that determine the use of language tests (Jenkins 2007; McNamara and Roever 2006). In tests, for instance, the function of gatekeepers is to set the parameters that serve a purpose in society, such as a condition for promotion, employment, citizenship, or immigration. Such gatekeeping practices include agendas that determine what standard of language allows an individual to practice medicine or to hold an academic position. These agendas are closely linked to ideologies of standard language and prescriptivism that position the NS as the only official authority on which to base decisions, a point further discussed in 3.5. In other words, the gates are controlled by the language ideologies that inform the gatekeepers who make the decisions about what is correct, appropriate, and proper language for a given context and situation (Jenkins 2007). Gatekeepers include influential groups, such as government institutions, language examination boards, publishers, universities, and other authorities.

Moreover, the policy-driven agendas of gatekeepers are common in public educational settings, where government officials impose language standards and where language examination boards determine the standard for NNSs entering a university. In Europe, an example of control is CEFR (Europe 2001), which determines the standards for language teaching, learning, and assessment throughout Europe. According to Jenkins (2007: 239), such gatekeeping practices influence the beliefs and attitudes of non-gatekeepers, who simply conform to the policy-driven decisions. The ELF literature also indicates that the current policy-driven standards based on standard language and NS language ideologies do not reflect the target needs of SLUs of English (e.g. Jenkins 2000, 2006, 2007). While many ELF users are misguided by the dominant language ideologies of nativeness and Standard English (Jenkins 2006: 143), some recognize the legitimacy of ELF and question the relevance of NS language ideologies.

### **3.4 My definition of language ideologies**

Taking the literature review from the previous sections into consideration, I will now outline my working definition of language ideologies. For the analytical purposes of this dissertation, language ideologies

- Can originate in an implicit assumption or an explicitly stated belief
- Are grounded in social experience where histories guide the beliefs
- Are situated in given socio-cultural contexts
- Are subject to power relations and dominance, and are instrumental in creating such relations

- Are held by individuals and may be shared by subgroups of individuals rather than by all members of a given speech community
- Exist in a heterogeneous speech community where multiple language ideologies co-exist and are potentially conflicting
- Inform social and linguistic practices, such as ascribing value to a certain linguistic feature that serves as an identity marker that may carry negative consequences, or index an individual into a specified group, or create hierarchies

From these points, it should be evident that I am approaching language ideologies from a critical view as opposed to a neutral one. To briefly summarize my position, language ideologies consist of socio-political notions of what constitutes language and what its social functions are within a given social group. These notions can be implicit and/or explicit, and although they constitute a shared system of beliefs that inform the way in which individuals interact with language, these positions are not without disputes as they represent multiplicity in different social contexts.

### **3.5 Some widespread language ideologies**

As discussed in the previous sections, the study of language ideologies bridges language use with social theory, linking conceptualizations of language use, attitudes and beliefs to discourse related to power and social differentiation. These links can be found at both macro- and micro-levels (Fairclough 1995). The primary ideological position suggests a monolingualizing tendency in an imagined political community (cf. Anderson 1983), where language is homogeneous. Such a position ignores the complexities of real, heterogeneous, multilingual societies, where variation in language is common and language ideological beliefs are not uniform. Moreover, the presence of a dominant language ideology of monolingualism raises questions about social justice in a heterogeneous, multilingual society. Similarly, a language assessment scheme based on language ideologies representative of a monolingual, homogeneous target culture raises questions about social justice for plurilingua-cultural ELF users.

In order to view language ideologies as socioculturally situated, it is necessary to investigate them in connection with the local context. In the present study, this means investigating the language ideologies of EMI lecturers, who are plurilingua-cultural users of English in an international academic context. Gaining knowledge about their language ideologies is important for the development of appropriate language assessment criteria for their target use of English in a multilingua-cultural environment that is outside the monolingual NS world. For this reason, this section covers four widespread language ideologies that will serve as the basis for analyzing the data in this study. It first presents language ideologies typical to monolingual, homogeneous societies before moving to those connected to multilingual, heterogeneous societies.

### 3.5.1 One-nation-one-language ideology

One language ideology that has been influential in the creation of a monolingual, homogeneous society is the *one-nation-one-language ideology*. This language ideology has played an important role in the formation of a nation state (e.g. Woolard 1998: 16). As its name implies, this ideology defines language through national as well as ethnic boundaries, which do not necessarily coincide. This language ideology equates one language with one culture or ethnic group, which in effect uses language to unite a community or nation. Although this language ideology played an influential role in 19th century Europe for creating statehood (Gal and Irvine 1995: 968), it is still widespread and particularly important in relation to minority languages maintaining autonomy (e.g. Trudgill 1974). This language ideology illustrates that political and cultural factors can be more important determiners than linguistic criteria, such as mutual intelligibility, for defining a language for a particular group. It also encourages monolingualism and may even stigmatize behaviors that belong to bi- and multilingualism, such as code switching or mixing. In short, a central goal of this language ideology is to define identity through the local national language and culture. Because a language ideology demarcated by national boundaries with one language presupposes a nation of NSs, this ideological position clearly is not representative of a group of lingua franca speakers communicating in a global community where national boundaries are irrelevant and where there may be no NSs.

### 3.5.2 Standard language ideologies

Monolingualism is also reinforced in a nation-state through a widespread language ideology, known as standard language. According to the Milroys (1999), the goal of standard language ideology is to promote uniformity in the structure of language use.

In the literature, discussions on Standard English are prominent, emphasizing its sociopolitical basis and its impact on language use in society (e.g. Cameron 1995; Holborrow 1999; Kachru 1997; Lippi-Green 1997; Milroy and Milroy 1999; Quirk 1990; Widdowson 1994). The following definition of *standard language ideology* is provided by L. Milroy

[A] particular set of beliefs about language ... . [which] are typically held by populations of economically developed nations where processes of standardization have operated over a considerable time to produce an abstract set of norms – lexical, grammatical and ... phonological – popularly described as constituting a standard language.

(1999: 173)

This definition draws on the historical dimension of standardizing a language, which signals norming as having an important role. This role is prevalent in the conformity discourse related to immigrants in the United States. For example, in the U.S., immigrants

are expected to acculturate to American norms, which means replacing foreign traits with characteristics that appear to be more ‘American’, including characteristics related to spoken and written English (Dicker 1996). This view is indicative of the official-English movement, based on the ideology that immigrants must conform to American language and culture in order to be accepted as part of their new country. This example illustrates the dominant ideology of monolingualism in the United States, where, “it is not ‘normal’ to speak a language other than English, nor is it ‘normal’ that, if you do, you would want to continue to speak it after having learned English” (Shannon 1999: 183). It also illustrates how the official-English movement represents a question about political identity in terms of who is American. This question of who is and who is not American can also be linked to another language ideology related to the images of ‘self’ or ‘other’, or ideologies of us versus them. While the ideology of monolingualism dominates in the United States, it is not uncontested.

In the previous example, the ideologies of norming toward a single standard variety shows how closely connected language is to social and political values as well as attitudes that determine how different dialects and accents are perceived and used to create indexical social categories. It is these motivations that gives impetus to a standard variety as having more status and prestige than other varieties (Trudgill 1974: 19). This position is also illustrated in the following comment where Vestergaard points out “This [widely accepted] variant is known as the standard variant, and this is also the variant that native speakers find the most suitable for foreigners, whether they themselves use it or not” (1996: 114). Vestergaard’s point illustrates the position of Standard English as the English language, where other varieties are marked as less desirable. In this way, the standard variety is utilized to persuade NSs and others that they cannot speak English (Trudgill 1974: 20). Simultaneously, it supports the official or standard variety in being a symbol of domination, which corresponds to the Gramscian (1971) notion of hegemony.

Several studies have shown that a standard variety can be misunderstood as representing greater moral or intellectual worth as opposed to contesting other varieties (Heller 1999; Jaffe 1999; Schieffelin and Doucet 1998; Watts 1999). As a result, speakers of official varieties may be viewed as having superior moral and intellectual value than those who speak unofficial varieties. This language ideology subjects speakers of unofficial varieties to symbolic domination and power that may be misrecognized as legitimate power (Bourdieu 1991: 170). In an orientation towards a standard variety that is masked as equal opportunity in a democratic society, individuals strive for homogeneity. In the process, they undervalue or overlook linguistic diversity and then pay the consequences of marginalization or penalization for non-conformity (Blackledge 2000: 28).

In defining standard language, the Milroys also describe it as entailing the belief in “one and only one correct spoken form of the language, modeled on a single correct written form” (Milroy and Milroy 1999: 174). This belief regulates language use, ensuring correct usage. They further point out that not adhering to standard language ideology carries a serious social consequence where “language change equates to language decay, and variation with ‘bad’ or ‘inadequate’ language” (ibid: 175). In this way, it ascribes a prestige value to the standard language ideology while simultaneously devaluing non-

standard varieties, where non-prestigious forms may be regarded as “‘wrong’, ‘ugly’, ‘corrupt’ or ‘lazy’” (Trudgill 1974: 20). These social values are based on prescriptive views of language use that entail judgments about ‘correct’ and ‘incorrect’ use of language, which are imposed by authority (Milroy and Milroy 1999). However, as the Milroys argue, such rules are just as arbitrary as dinner table etiquette. For instance, in Europe the fork is in the left hand while the knife is in the right hand (at all times), and in the USA the fork in the right hand except when the knife is needed, and any deviation from the rule is ‘bad manners’ (1999: 1). Thus, prescriptivism can include pronouncements of what a particular group considers to be ‘good taste’.

Thus, the role of prescriptivism relates to language standardization, where rules govern how language is used. Such rules are based on written language grammar and acceptability rules. These rules can cover standards for spelling and grammar as well as what rules for what are socially correct. Because a language standard is primarily maintained through the writing system, the written word becomes the model of correctness. Thus, the writing system serves as the source of prescriptive norms. It is motivated by political, social, and commercial needs, and imposed on the norms of language usage through language policy and authorities who codify the language in dictionaries and other similar references, as well as through public channels such as the education system. The ideology of standardization views language as a relatively fixed and invariant entity. For instance, spelling is the most uniform element of language use, and it contrasts with the variability in spoken language.

A central problem with prescriptivism is that it is based on a rather narrow definition of grammar, which aims to develop norms for clear, unambiguous written prose (in Standard English) and many of which are inappropriate for spoken language (Milroy and Milroy 1999). Whereas the Milroys discuss prescriptivism from the viewpoint of society more widely, that is, from the view of ordinary people, Linell targets linguistics and presents a similar view:

Our conception of language is deeply influenced by a long tradition of analyzing only written language, and ... modern linguistic theory, including psycholinguistics and sociolinguistics, approaches the structures and mechanisms of spoken language with a conceptual apparatus, which ... turns out to be more apt for written language in surprisingly many and fundamental aspects. I ... refer to this situation as the written language bias in linguistics.

(Linell 2005: 1)

What these arguments signify is the dominance of the prescriptive, written word as the basis of correctness for spoken and written language. Thus, utterances in speech claimed as ungrammatical or inappropriate are most likely based on prescriptive arguments. This results in the norms of written prose being misapplied to spoken language. When spoken language is grammatically judged in terms of standard written norms, it results in reducing variability in speech, limiting conversational styles and appropriateness in varying situational contexts (ibid). Furthermore, in some spoken contexts literary grammar may be inappropriate and non-functional.

Moreover, mode is one of the most obvious differences between spoken and written language. Speaking, for example, occurs in a situational context (i.e. context-tied), whereas written language is context-free. Because spoken language is context-tied, it is characterized by inexplicitness (Sinclair quoted in Warren 2006), situational ellipsis and deictic expressions that become unspecific when context-free, as well as vagueness in vocabulary (Mauranen 2007b lecture series on spoken language and ELF) whereas written language requires being more explicit, using language that carefully specifies the objects, persons, events, and actions (Milroy and Milroy 1999). Even though the nature of spoken language is very different from written language, prescriptivism makes no allowance for variation in the spoken mode. A standard ideology promotes uniformity at the expense of variety, and the prescriptive tradition has always aimed at uniformity in writing as well as in speech. Although attempts have been made to prescribe uniformity in speech, it has not been very successful. As Trudgill points out “[t]here is no universally acknowledged standard accent for English, and it is in theory, possible to speak Standard English with any regional or social accent“ (1974: 19). He further states that the only exception to this point is British English: accents in Britain are tied to Standard English (ibid). The lack of success in standardizing spoken English is partly attributed to the failure to consider the extent to which variability in spoken language results from social factors and speech functions (Milroy and Milroy 1999).

As a regulator of language use, prescriptivism began at a time when the means for studying spoken language in very much detail were limited. More recently, modern technology has produced many recording devices, such as tape recorders, iPods, and mp3 players, all of which provide a way to study specimens of spoken language, and it has become clear that there are great differences between spoken and written language. Thus, it is conceivable that the role of prescriptivism in spoken language may recede, particularly in contexts not dominated by NSs of English, as modern techniques for studying spoken language and describing it are now thriving. Shifting from prescriptivism to native speakerism, the next section discusses L2 learner-related language ideologies.

### **3.5.3 Native speaker language ideologies**

In section 3.2, three standard language ideologies were introduced along with three corresponding ideological processes (i.e. iconization, recursiveness, and erasure). This section extends the discussion to include NS learner-related language ideologies. A review of the SLA literature puts the concept of the NS at center stage. In SLA research, the primary function of the NS is to serve as the benchmark for SLLs, even though efforts to define it have proven to be problematic (see Davies 1991, 2003). By taking the NS as the benchmark, the ideology of standard language is extended to include a learner-related language ideology, known as *NS language ideology*. The following classic definition by Crystal introduces a working idea of some central tenets that give impetus to learner-related language ideologies:



A term used in LINGUISTICS to refer to someone for whom a particular LANGUAGE is a 'native language' (also called 'first language', 'mother tongue'). The implication is that this language, having been acquired naturally during childhood, is the one about which a speaker will have the most reliable INTUITIONS, and whose judgments about the way the language is used can therefore be trusted. In investigating a language, accordingly, one is wise to try to obtain information from 'native-speaking' informants, rather than from those who may have learned it as a SECOND or foreign language (even if they are highly proficient). Many people do, however, develop a 'native-like' command of a foreign language, and in BILINGUALISM one has the case of someone who has a native command of two languages.

(Crystal 1985: 204, emphasis in original)

For example, this definition designates the NS as the only authority on questions of language use, where all NSs are deemed as having the same 'perfect' language abilities. By signaling the NS as the only trustworthy source, this definition promotes an ideology of the NS having superior linguistic expertise and knowledge of the language that translates into the NS as expert. This ideology is based on the assumption of correctness of NS language, an idealized view of NSs in a homogeneous, monolingual speech community that ignores the reality of broad variation present among the L1 abilities of its speakers. Present in the definition is also the assumption of NS ownership of the language, a belief that supports the one-nation-one-language ideology as well as standard language norms. The above definition also includes an ideology of 'othering' by making a clear distinction between NSs and NNSs, an us-versus-them ideology, which applies regardless of the proficiency levels (of either). Implicit in this 'othering' ideology is the belief that all NNS language is inferior or deficient to NS language, in one way or another. This definition also illustrates the notion of the native speaker construct in theoretical linguistics, where it represents an "ideal speaker-listener, in a completely homogeneous speech community" as defined by Chomsky (1965).

Some scholars, however, argue that this idealized NS model creates a "monolingual bias in second language acquisition research" (Cook 1997: 35) and view the NNS as "a defective communicator, limited by an underdeveloped communicative competence" (Firth and Wagner 1997: 285). Others view the NS as more of a myth that "exists only as a figment of the linguist's imagination" (Paikeday 1985: 25) than a reality (see also Davies 2003). Seidlhofer (2003) takes the discussion further emphasizing "realistic" English over "real" English. She argues that "native-speaker language use is just one kind of reality, and not necessarily the relevant one for lingua franca contexts" (Seidlhofer 2000: 54). She further claims that there is no justification for calling an item an error "if the vast majority of the world's L2 English speakers produce and understand it" (ibid: 65). On the NS as a model, Seidlhofer (2000) advocates abandoning it, a position that echoes Kramsch who earlier suggested that we should "take our cues not from monolingual native speakers ... but from the multilingual non-native speakers that constitute the majority of human beings on the planet" (1993: 49). Seidlhofer (2000) views the NS model as not appropriate now

that the number of NNSs outnumbers the ENL speakers, bringing its usefulness as a model into question. She also points out that the native speaker is an icon widely deemed as linguistic imperialism in the literature (e.g. Canagarajah 1999; Medgyes 1994; Paikeday 1985; Phillipson 1992).

Discourse related to NS language ideologies also includes native speakerism, a term coined by Adrian Holliday<sup>9</sup>. He uses the term to capture the belief in native speakers being better qualified to teach language (with or without teaching credentials) than non-natives. The discourse includes assumptions that NSs are authorities on their language and have superior competence, that L2 speakers must treat NS competence as the target, and that NSs are best for teaching language. Supporters of this inherent NS superiority include Prator (1968) and Quirk (1990), who treat varieties spoken by multilingual NNSs as interlanguages that approximate the NS norm. Other scholars, however, argue that NS superiority is anachronistic and goes against linguistic relativism, which postulates that no status differences exist between languages based on purely linguistic grounds, even though extra-linguistic reasons also exist (see Braine 1999; Canagarajah 1999; Kachru 1991). What native speakerism also misses is that languages in situations of contact will undergo changes, including modes of appropriation.

NS language ideologies have an impact on professional life, from employment policy to 'othering'. They separate NSs from NNSs, not only through language but also through cultural bias. An 'othering' theme can include a range of confining and negative labels, such as 'non-native', 'collectivist', 'passive', 'undemocratic', or 'hierarchical' (see Holliday 2005). In addition, employment policy may block non-native speakers with a single clause 'native speaker only'. What such policies fail to consider is that language proficiency is only one part of what may make a good employee. Having a good basis in language does not necessarily entail having the particular skills and knowledge needed for a given job. For example, a NNS who is skilled in teaching will be more effective than a NS who is not. A study on ELT teachers that supports this view cited factors such as "being qualified", "being prepared", and "being professional" as more important for an instructor than being a native speaker (Braine 2005: 19). Such studies make a valid point that should be considered in the employment of EMI lecturers, for example.

As a learner-related ideology, the native speaker upholds political and cultural beliefs related to norming as supported by a one-nation-one-language ideology, where maintaining this ideology appears to be more important to linguistic criteria than realistic criteria relevant to measuring L2 performance. Like the previous ideologies introduced, this language ideology encourages normative, prescriptive goals for supporting monolingualism. Upholding the belief in the idealized NS results in an automatic rejection of other standards or varieties. Thus, L2 speakers who fail to achieve native-like proficiency are viewed as having inferior language, which includes the majority of English speakers. In other words, it subjects SLLs and SLUs to normative targets that are unattainable by the majority, and SLUs to linguistic criteria that are irrelevant to ELF communication in international contexts, a point discussed in the next section.

---

<sup>9</sup> Adrian Holliday introduces the term 'native speakerism' in his (2005) book, *The Struggle to Teach English as an International Language*.

### 3.5.4 English-as-the-global-language ideologies

The aforementioned language ideologies are representative of colonial and post-colonial times. In a post-modern world, however, these language ideologies seem fossilized as they do not capture the beliefs present in new and emerging communities of practice among ELF users in a flattened, internationalized world (cf. Friedman 2005). By flat, I am implying the absence of language ideologies that are indexical or hierarchical. Among these speakers, a widespread language ideology is the belief in *English as the global language*. This ideology is discussed in Haberland (2009) in terms of hegemonic discourse, where English is viewed as the only language relevant to international communication. The hegemonic discourse relates to the general consent of speakers of different languages choosing English for communicating in a vast number of situations because they “consider this choice natural with respect to the existing linguistic world order” (ibid: 25). Their choice relates to the use of English in ‘the world at large’ and lacks an integrative motivation to a specifically defined target language community. This ideology is present among speakers with multilingua-cultural backgrounds in a heterogeneous international context, who choose ELF for communication.

ELF can also be seen as belonging to a Global Englishes paradigm, where all Englishes are equally accepted, rather than assessed against ENL norms (Jenkins et al. 2011) Supporting the view, Jenkins argues that ELF is “freed from the STANDARDIZING CONSTRAINTS” (2011: 291, original emphasis). As Jenkins (2006a; 2009) contends, this does not mean that there are no standards. Taking these views as a baseline, we would expect that ELF users negotiate some rules of linguistic appropriateness, and especially in the absence of some imposed authority. A study lending support to this notion is Hynninen (2013). She employed a bottom-up approach to the study of language regulatory practices among ELF users, where she investigated the construction of living norms in academic spoken discourse. She found that ELF users draw on ENL norms but also employ regulatory practices as a means to enhance mutual understanding through mediation, embedded repairs, language commenting and language correcting. Her findings also show that correction practices in ELF differ from NS-NNS interaction (2013: 252). This research illustrates a belief in ELF speakers as ‘users’ of English in their own right, where ENL norms were not automatically relevant as the measure of acceptable English. It also supports the notion of assessing the quality of English in a given community against the practices of that community (Canagarajah 2006) as opposed to how closely it approximates an ENL variety.

English-as-a-global language ideologies include the notion of English as a language for communication as opposed to a language for identification. This distinction is useful: English as a language for communication embraces the notion of English being the ‘property’ of all its speakers, who have native languages that will continue to function as languages of identification. In other words, their L1 identities serve as a means to connect with the cultural heritage of their L1 community (see Jenkins 2007). This distinction also highlights that English is not one monolithic, hegemonic voice, but a diversity of different voices. Users of English perceive its use as a tool, where it serves as a medium for communication that gains substance from the different national, regional, and individual

cultural identities that ‘users’ bring to it (House 2001). In other words, English itself does not bear national identities for SLUs because they do not see it as a language for identification. Thus, English unites SLUs, who view it as owned by all and as a means for enabling communication as opposed to being a cultural symbol.

Because of its spread around the world, English is also viewed as “de-nativized” to an extent (House 2001). Arguments for this view include the diversification of English into many NNS varieties through nativization processes (see Braine 1999; Singh 1998). A de-nativized view is also supported by the huge number of non-native users, which is substantially larger than its native speakers (ibid). Nor is English viewed as being NS owned (see Widdowson 1994).

### **3.6 Conclusion**

The theoretical framework presented in this chapter will serve as the point of departure in seeking language ideologies held by my research informants, the analysis of which is carried out in Chapter 7. From the present chapter, the central point of the discussion can be summed up as follows. Language ideologies enter speakers’ perceptions of how they fit into local or global contexts. In this way, language ideologies serve to either differentiate or to unify individuals. The picture, however, is rather complicated as relationships in post-modern societies have become increasingly complex in a world that has grown smaller. With the widespread use of English, the traditional monolingual language ideologies that promote homogeneity, such as standard language and one-nation-one-language ideologies, do not capture the beliefs held about the current linguistic world order, where speakers are multilingua-cultural and form heterogeneous social communities in a flattened world.

## 4 Assessing spoken English: ELF, LSP, and self-assessment

This chapter situates the present study in L2 language assessment research related to general language proficiency and professional skills. More specifically, the chapter draws attention to concerns related to the language assessment of SLUs of English. Because today's global community includes a large number of ELF users in a variety of academic and professional contexts, the traditional approach to language assessment that involves distinguishing between *us* (i.e. NSs) and *them* (i.e. NNSs) seems off-target in relation to the present needs of all parties involved. This point is particularly important as language assessments serve the purpose of distributing coveted resources, such as jobs in specialized fields, where the question of how to distribute such resources is often based on ideological views (as discussed in Chapter 3) and are policy-driven (as discussed in Chapter 2).

Thus, this chapter explores the concept of language assessment from different perspectives. First, to gain some insight into performance assessment, two current traditions are reviewed from the viewpoint of capturing the social aspects relevant to performance before examining how closely their related assessment criteria meet the challenges of real-world needs. Second, one might wonder how well current assessment criteria meet the needs of ELF users. To address this question, this section compares two proficiency paradigms and presents a rough checklist for determining the broad needs of SLLs versus SLUs. The two remaining sections cover literature on professional assessment relevant to this study, including language for specific purposes and self-assessment.

### 4.1 Performance assessment

The primary purpose for reviewing performance assessment is to grasp an understanding of the most common traditions in measuring L2 performance and to establish how closely such traditions represent the real world needs of the interested parties. One area of interest is the extent to which current traditions capture the social dimension in performance assessment. This factor is particularly relevant to the development of language assessment criteria for ELF users, who emphasize communicative efficiency over linguistic accuracy. To achieve this goal, this section first briefly touches on the meaning of performance before presenting the two most common types of performance assessment in L2 testing, and then highlights some key aspects relevant to the development of assessment criteria.

This review begins with a general definition: a performance test is “one in which some criterion situation is simulated to a much greater degree than is represented by the usual paper-and-pencil test” (Fritzpatrik and Morrison 1971: 238). According to this definition, simulation is one criterion of performance assessment. This criterion signals a relationship between the test task and reality, where the aim is to simulate a relationship that is direct,

close, and authentic (McNamara 1996: 11). In addition, this relationship depends on two factors: the degree to which the simulation captures reality, and the extent to which the simulated performance is relevant to the criterion (real-life) situation (Fritzpatrick and Morrison 1971: 240). Not only should the performance be representative of a real-life performance, it should also bear a relationship to valued performances in the real world (ibid). Together these points characterize a performance test as relatively realistic.

Central to the discussion on modern-day language assessment are two concepts: *knowledge* and *performance*. While the concept of language knowledge is fairly straightforward, the concept of performance is not. *Language knowledge* refers to knowledge of formal linguistic rules as related to grammatical rules, sociolinguistic rules and procedural rules. This definition extends beyond Chomsky's (1965) notion of underlying *language competence* to include Hymes' (1972b) notion of language knowledge as including sociolinguistic rules. On the concept of *performance*, while Chomsky viewed language as actual instances of use, Hymes saw ambiguity in the term. He thus proposed the following distinction: *actual use of language* (in real time; also known as actual performance), and *ability for use*, which refers to the underlying potential of a speaker to realize a performance (ibid). Hymes' dichotomy is helpful in understanding discussions related to communicative language testing and performance-based testing. In communicative language testing, the test construct generally relates to *language competence*, also known as *language proficiency*<sup>10</sup> or *language ability*, where these terms include language knowledge and performance, but exclude most non-linguistic factors relevant to ability for use. In contrast, in a performance-based test, the test construct may be grounded in *ability for use* rather than language competence, where non-linguistic factors generally carry heavier merit than linguistic ones (McNamara 1996). In addition to Hymes' distinction in the term 'performance', there can also be other meanings of it in test specifications, a point discussed in the next section.

A review of the literature indicates that there is no unified theory of language proficiency (e.g. McNamara 1996, Bachman 1990). As a term, it can be used to refer to language knowledge, competence, and ability – regardless of the conditions under which language has been learned (see, for example, Carroll 1961, Davies 1968, Spolsky 1968, Oller 1979, Bachman 1990). In discussion, I use the terms language proficiency and language competence interchangeably to refer to the ability of an individual to speak and perform in an acquired language. However, in Chapter 5, I discuss language competence as operationalized in CEFR<sup>11</sup>, where the term seems to be closely linked to an SLA paradigm, based on Chomskyan linguistics.

---

<sup>10</sup> In the literature, the term 'language proficiency' can be used interchangeably with language competence and language ability or then defined to take on more specific meanings by an author or test construct.

<sup>11</sup> The CEFR provides an operational definition of language competence through its descriptor scales (discussed in Chapter 5) and uses the term language proficiency more generally.

#### 4.1.1 Performance in second language assessment: two traditions

To understand the theoretical basis of a performance assessment in second language testing, it is necessary to note the different meanings of the term *performance*, which can be interpreted through test specifications as well as through practices. McNamara (1996: 26) outlines three common ways in which performance may be used. First, from a *theoretical position*, performance may be used to signal underlying language ability, as described in the modern-day works of testing experts such as Bachman and Palmer (1990; 1996). Another use of the term includes *skilled execution*, where the focus is on displaying a skill, such as a physical performance of gymnastics. A third use is to signal *performance on real-world tasks*, where the assessment precludes a direct simulation of a real-world situation. However, broader definitions are also possible. On this point, McNamara argues that performance testing in L2 assessment features one simple requirement: the assessment occurs while the candidate is “in the act of communication” (cf. Savignon 1972, cited in McNamara 1996: 26). This broader view is the one adopted in this dissertation.

In L2 assessment, performance testing can be traced to the 1960s. The need for this type of assessment arose primarily from two developments. The first pertains to a practical need, that is, a selection procedure for foreign students to English-taught programs in tertiary education (McNamara 1996: 25). The second need concerns addressing theoretical advancements related to language teaching. In other words, the need to align language assessment with evolving language teaching practices, such as the communicative competence movement (ibid). Against this background, two traditions to L2 performance assessment developed: the work-sample tradition and the cognitive-psycholinguistic tradition (ibid). Here, I will briefly highlight some key aspects of each.

McNamara (1996) describes the work-sample tradition as originating in the work of Carroll (1961 [1972]) who proposed developing a test that would assess the *ability to function* in a target language, regardless of the L1. In other words, the test purpose is to predict performance in a real-world context. From Carroll’s proposal, two assessment methods developed: *the work-sample method* and *the simulation method*. The main difference in the two methods is that the performance assessment in the work-sample method occurs on the job at the workplace while the simulation method occurs in a simulated environment where tasks are abstracted from a workplace reality (McNamara 1996: 14-15). Both methods assess whether an individual has the ability to perform tasks associated with a particular job, and both methods are used in assessing teacher performance (e.g. Elder 1993; 1994). According to Jones (1985: 19), the simulated method is the most practical for language assessment.

Work-sample assessment is based on domain sampling that takes a *task-centered* approach to performance assessment, as differentiated from a *construct-centered* approach based on a theory of language ability (McNamara 1996: 6). In the words of Messick (1994), the *target* in the work-sample assessment is to assess performance, where the *medium* of performance is a second language. Although this approach specifies the sociolinguistic context in accordance with Hymes’ concept of sociolinguistic competence (1967; 1972b) that includes an *ability for use*, it is viewed as pragmatic and atheoretical (McNamara 1996: 25). It is also “characterized as behavior-based and sociolinguistic in

orientation” (ibid). While the approach may be weak in terms of language assessment theory, it is strong in its criteria that are primarily based on real-world criteria as needed for the fulfillment of the tasks (ibid). Because of its sociolinguistic orientation, work-sample assessment offers an appealing base for the development of performance-based assessment criteria for SLUs who typically use English in academic and professional domains to accomplish work-related tasks.

Unlike the work-sample tradition, the target of the cognitive-psycholinguistic tradition is primarily to assess what a performance sample reveals about the underlying language ability or knowledge<sup>12</sup>, the construct to be assessed<sup>13</sup> (McNamara 1996: 6). From the point of view of Messick (1994), this type of performance could be seen as the vehicle of the assessment. In second language assessment, this tradition to performance testing began with the work of Lado (1961) in a structuralist framework, which attempted to capture language knowledge. Although the cognitive-psycholinguistic tradition has its roots in pre-communicative assessment, it carried the measurement of language knowledge over into assessments that were communicatively oriented (for an early example, see Savignon 1972). In modern-day communicative language assessment, this tradition links to the work of testing experts, such as Savignon (1972; 1983), Canale and Swain (1980), Canale (1983; 1984), and Bachman (1990), who have advanced the theoretical foundation of the communicative approach. The advent of the communicative movement also sparked a debate on authenticity. In language testing, this debate first focused on whether test tasks should mirror ‘real world’ tasks, and whether that was actually possible (Morrow 1979). Since then, authenticity in language testing has progressed to include test-taker involvement in test tasks (see Bachman and Palmer 1996 for details of the theoretical framework).

However, capturing authenticity in language testing continues to be a persistent unresolved issue (e.g. Leung and Lewkowicz 2006; McNamara 2001; McNamara and Roever 2006; Widdowson 1998a, 1998b). For example, the language proficiency literature

---

<sup>12</sup> Note that language proficiency can also be defined as language knowledge and/or as language ability. The term, language knowledge, generally refers to knowledge of formal linguistic rules, including grammar, lexicon, sociolinguistic correctness and other linguistic aspects. The term, language ability, is less straightforward since it involves questions of ability for use, which includes cognitive and affective factors (e.g. emotional state, personality factors, reasoning powers). Thus, a clear definition of language ability is essential to language test development and use. Such a definition is generally derived from a language-teaching syllabus or a general theory of language ability. Although many L2 language proficiency tests continue to be based on a skills framework, language testers nowadays take a broader view of language ability than that proposed by Lado (1961) and Carroll (1961). This broader view recognizes the importance of context beyond the level of sentence, thus giving consideration to the discourse of which the sentences form a part and the sociolinguistic context that governs the nature of the discourse (form and function). Thus, various definitions of language ability are derived from the test construct and the way it is defined operationally (e.g. through the rating scales).

<sup>13</sup> A test construct defines the psychological concepts to be measured in an assessment. In other words, it defines specifically the concepts about which a test (examiner) is trying to gather evidence that will be used to make a decision about the test taker.



criticizes the Bachman (1990) model, which is still widely used, for its inability to account for social aspects of language proficiency, such as interlocutor behavior (e.g. Lumley and Brown 1998; Luoma 2004), test-taker characteristics (e.g. O'Sullivan 2000), test-task familiarity (e.g. Foster and Skehan 1996; Wigglesworth 1997), and personality types (e.g. Berry 2004). In addition, Leung and Lewkowicz (2006) criticize the Bachman-Palmer (1996) model as being impractical for heterogeneous test populations (such as ELF users), pointing out the improbability of being able to define the characteristics of the target language use domain. McNamara (1996: 75) also criticizes these models as being focused on measuring individual traits or qualities, resulting in a one-sided relationship. He argues that a view of communication as jointly constructed would be more relevant. In this way, current models of language proficiency are still weak on capturing what Hymes referred to as *ability for use*. For the purpose of developing assessment criteria for SLU, such models appear to be too limited in their ability to capture communicative factors relevant to real-world communication.

Although these two different traditions, work-sample and cognitive-psycholinguistic, have developed in L2 performance assessment, they have some commonalities. One common point is that they are designed to elicit performances with specific language behaviors that testers wish to assess. Even though performance assessments are designed to elicit underlying language ability of an individual, they are typically rated in terms of the linguistic characteristics of the performance that the test designer considers to be important and relevant. Another commonality is in their focus on the importance of building theoretical frameworks that reflect performance assessments needing to correspond to real-life roles, tasks, and situations. A major difference, however, is in their orientation to assessment criteria, the topic of section 4.1.2.

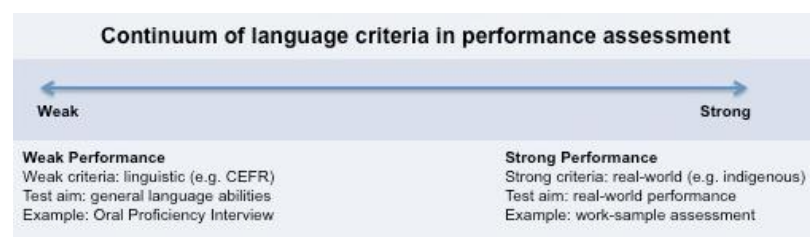
#### **4.1.2 Performance-Based Assessment Criteria**

As indicated in the previous section, current cognitive-psycholinguistic models of language assessment are designed to measure traits or qualities that reside in an individual. This feature is an outcome of psychometric testing, which is common in L2 performance assessment. The strength in psychometrics is in its ability to provide a framework for accurate test instruments that assess the psychological construct under investigation (McNamara 1996). Thus, its appeal is in its rigorous methods, which provide evidence for justifying inferences (ibid) and which serve to reduce bias and unwanted impact (McNamara and Roever 2006). However, this approach to language assessment minimizes the social dimension, which may be deliberate (ibid). Its treatment of social context is used to project a set of demands onto learner language ability (ibid: 32). Such approaches to language assessment illustrate working from within the test itself where there is no “theory of the social context in its own right” (ibid). This lack of a social theory in performance assessment is problematic, and generally results in assessment criteria that are heavily linguistic in orientation. The question, then, is whether or to what extent this approach meets the real world needs of test-takers and other relevant parties.

On this point, McNamara (1989; 1996) captures the essence of these two performance assessment traditions in a hypothesis of weak to strong performance, based on the distinction between ability to do future tasks and ability to use language in future tasks. The former type of inference includes non-language factors, such as personality characteristics, while the latter excludes such factors. Using this distinction, he proposes that assessments strong on testing are weak on language and vice versa (McNamara 1996). McNamara contends that an assessment strong on testing leads to reliable tests with results that can be generalized, but which may be ill-fitted to the needs of the language user. Conversely, an assessment strong on performance that makes relevant claims about the success (or lack of it) in completing the targeted task leads to results that lack psychometric qualities and consequently are less reliable (ibid). His dichotomy places language performance assessments into two broad categories that form the endpoints of a continuum (ibid: 43). On this continuum, language performance assessments range from weak to strong, based on the degree to which their language assessment criteria approximate real-world criteria. Figure 4-1 illustrates this concept.

This continuum places real-world criteria, for example used in work-sample assessments, at the strong end and assessments that are entirely based on linguistic criteria and aimed at general language abilities at the weak end. Examples of assessments

**Figure 4-1. Continuum of language criteria in performance assessment, based on McNamara 1996.**



that use weak criteria include the Oral Proficiency Interview and the CEFR.

While an assessment using weak criteria may indicate more about the basic language skills in terms of accuracy and appropriateness in relation to a standardized norm, an inherent weakness in such scales is in their inability to measure ability for use, which includes a number of non-linguistic factors. This point is perhaps clearest in the words of Jones in his discussion on the role and implications of non-linguistic factors in performance-based language assessments:

With regard to second language performance testing, it must be kept in mind that language is only one of several factors being evaluated. The overall criterion is the successful completion of a task in which the use of language is essential. A performance test is more than a basic proficiency test of communicative competence in that it is related to some kind of performance task. It is entirely possible for some examinees to compensate for low language proficiency by astuteness in other areas. For example, certain personality traits can assist examinees in scoring high on interpersonal tasks, even though their proficiency in the language may be substandard. On the other hand, examinees who demonstrate high general language proficiency may not score well on a performance because of deficiencies in other areas. (Jones 1985:20)

His statement draws attention to the main objective in assessing a performance: to measure successful completion of a task, which requires giving consideration to both linguistic and non-linguistic factors relevant to real-world criteria, where non-linguistic factors may even receive more weight than linguistic ones in assessing a performance.

Not only are assessment criteria weak to strong, they also represent valid measures of the test construct. This point is recognized in Messick's (1989) validity theory. Although his theory represents an individualist, psychological approach to measurement that is concerned with fairness, it also incorporates a social dimension that addresses test use, as shown in Table 4-1.

On the social dimension, his theory argues two main points: 1) the notions of what we measure and what we focus on in measurement will reflect values, which can be assumed to be rooted in sociocultural origins, and 2) test results effect the social contexts in which the scores are used, and this matter needs to be a concern of those in charge of an assessment (ibid).

**Table 4-1. Facets of validity from Messick 1989: 20.**

	Test interpretation	Test use
Evidential basis	Construct validity	Construct validity + Relevance / utility
Consequential basis	Value implications	Social consequences

While hugely influential and valuable, Messick's work has some limitations, particularly in terms of the social dimension of language assessment. His theory carries forward the psychological tradition of measurement, similar to Cronbach (McNamara and Roever 2006: 12). Like Cronbach, Messick perceived assessment as inference-based reasoning and evidence gathering, where meaningfulness and defensibility were primary (ibid). For instance, tests and their assessments, such as IELTS and TOEFL, involve making inferences about how a test taker will perform in non-test settings. The assessment scheme is modeled on the demands of a target setting and predicts the performance of an individual in relation to this construct. This test scheme is inherently subjective as it involves matters of belief and opinion regarding both the test construct (e.g. what test designers believe is required in the relevant settings) and the evaluation of an individual's performance that is supported with reasoning and evidence (based on value-laden criteria), which is used to make a decision about the individual (McNamara and Roever 2006). This systematic approach constrains the inferential process where the aim is to guarantee fairness or validity, both to the test takers and to the receiving institutions. However, the commitment to working internally, within a test to develop test schemes, reflects the psychometric tradition upon which Messick's validity theory is built and which limits his theory to the technical qualities of assessments in terms of fairness (ibid). In this way, his influential work remains inadequate as a conceptual approach to understanding the social dimension of tests (ibid).

As Messick's theory rightly captures, measurements reflect values, some of which are sociocultural. This point is crystallized in a well-known L2 construct, that is, the native speaker. Arguments related to this value-laden concept also resonate in the debates on the notion of the native speaker as the standard measure for all Englishes (e.g. Jenkins 2007; Kachru 1986; Kachru 1992; Quirk 1990). The failure of cognitive-psycholinguistic

language assessment models to account adequately for the social dimension in performance-based assessment is likely to have some impact on assessing spoken professional English in international contexts. For SLUs, what is the value of rating scales that base an assessment on NS norms? Or a test that marginalizes the social context of use? Current assessment models that focus on measuring qualities or traits that reside within an individual are fundamentally asocial (McNamara and Roever 2006). This aspect presents a problem for performance assessment since language is closely interwoven into the tapestry of society where social and cultural interactions primarily take place face-to-face.

On value-based assessment, McNamara (2001: 339) raises a provocative question as to whether “the act of testing itself constructs the notion of language proficiency”. In other words, the test developers, raters, and rating scales construct the language proficiency rather than the traits that reside in a test taker. This provocative thought appears to represent a plausible reality. As Alderson and his colleagues (2005 [1995]) explain:

[E]very test has a theory behind it: some abstract belief of what language is, what language proficiency consists of, ... what language users do with language. This theory may be more or less explicit. Most test constructors would be surprised to hear that they have such a theory, but this does not mean that it is not there, only that it is implicit rather than articulated in metalanguage.

(Alderson et al. 2005 [1995]: 16)

This passage suggests that the beliefs of test developers construct the assessment scheme that forms the basis of language proficiency. This scheme includes rating scales with criteria that are applied to a test taker, the result of which could be seen as constructing language proficiency. Moreover, these rating scales and the resulting assessment outcomes are also treated as though they represent an objective reality in the real world. But whose reality? What is the communicative value of a scale that emphasizes correct usage of the definite article? For ELF users, the imperfect use of the definite article may suggest a lack of grammatical control of that particular form, but it does not necessarily hinder communication (Jenkins 2000). An examination of existing rating scales and assessment frameworks for international tests of English reveals an emphasis on NS norms, whether implicit or explicit, placing accuracy in lexicogrammatical features and correct pronunciation as central, especially for speaking and writing skills at an advanced level (see, for example, descriptors and bands for CPE, IELTS, TOEFL, TSE and any other rating scales).

For instance, a review of the spoken assessment criteria in the aforementioned scales also reveals accent to be one prescriptive measurement. This measurement is based on the degree to which a foreign accent is present, as determined by recognizable or noticeable L1 features in the L2. While accent can be an identifying feature, it does not necessarily measure intelligibility. It does, however, indicate a social preference for a prestige NS accent, which could be viewed as a form of accent discrimination, especially given that the majority of NSs fall outside this norm (e.g. Leung and Lewkowicz 2006, Jenkins 2000, 2007). Alternatively, accent could be seen as a positive feature, signaling a speaker who

wishes to maintain his/her L1 identity when speaking an L2 (Jenkins 2000). Another problematic area is fluency based on speech rate rather than continuity, which could also be seen as a NS bias (Levis 2006). Rather than individual-focused fluency criteria, Hüttner (2009) proposes fluent interaction as a measure of ELF fluency, for instance, through the application of Fillmore's (1979 [2000]) criteria, including continuity, coherence, context-sensitivity, and creativity.

Not only are NS norms a problem in L2 performance assessment, but also the categories NS and NNS. As distinctive categories, these terms have proven problematic, and have been questioned on both conceptual grounds (see, for example, Coulmas 1981 and Davies 1991, 2003) and empirical grounds. For example, studies of native speaker performance on EAP tests, such as the IELTS, have shown that results lack uniformity and that few NSs attain the status of expert user (see, for example, Clapham 1994, 1996; Lopes 1992). These findings suggest that reference to the NS as a benchmark in descriptor scales is not valid. Nevertheless, the categories of NS and NNS continue to represent a crude dichotomy that serves as the basis for many test constructs, while ignoring the fact that NNSs may be just as capable as NSs (McNamara 1996).

With the majority of English speakers in the world being non-native, real-life practices point towards the NS as having little relevance as a proficiency measure in international contexts. SLU of English is context-driven in a globalized society of lingua franca users, who determine their own norms (e.g. Jenkins 2007; Mauranen 2011a; Seidlhofer 2011). For ELF users, the literature suggests a strong need to adjust assessment practices to extend beyond the native speaker (e.g. Canagarajah 2005; Jenkins 2000; Jenkins 2007; Leung and Lewkowicz 2006; McNamara and Roever 2006). This situation calls for assessment criteria derived with the interests of ELF users in mind. This point is discussed in the next section.

## **4.2 Spoken ELF in assessing English performance**

Carrying forward the relevance of the social dimension in developing assessment criteria for the real-world needs of test takers, an understanding of ELF needs is central to identifying the qualities relevant to assessing their professional English. It is also a necessary step to prevent the act of testing from constructing language proficiency for ELF users through irrelevantly tuned criteria. Thus, this section presents a literature review on the language needs of SLU as contrasted with SLL.

The concept of analyzing SLL needs has been widely employed in language teaching since the 1970s, and analyzed in numerous works (e.g. Dudley-Evans and St. John 1998; Ellis 2003; Long 2005; Van Hest and Oud de Glas 1991; Vandermeeren 2005). A recent position of needs analysts also includes giving consideration to learner 'rights' in target choices (e.g. Benesch 1996, 1999, 2001). Jasso-Aguilar (2005) also argues for an approach to L2 rights when a needs analysis identifies a conflict between institutional practices and worker realities:

In conducting an NA [needs analysis], it is necessary to examine the social context in which the actors live their lives *critically* [italics mine], as well as the power differentials involved. As researchers, we must strive for a critical perspective based on dialogue with, rather than observation and manipulation of, people.

(Jasso-Aguilar 2005: 150)

His criticism is drawn from a triangulated needs analysis for the curriculum of hotel maids, where he served as an ESP consultant. This critical view positions itself with a sociolinguistic approach to language teaching and assessment. Moreover, critical views that focus on such power relations are increasing in ESP/EAP literature (e.g. Canagarajah 2002; Jenkins 2007).

Generally, a needs analysis targeting the objective needs of language use will address situational activities, competence targets, sociocultural contexts, relevant content (e.g. Dudley-Evans and St. John 1998; Richterich and Chancerel 1977) as well as identify competence gaps (e.g. Van Hest and Oud de Glas 1991). However, they fail to address the concept of identity in terms of cultural identities. For SLLs, learning involves becoming a member of a target culture of NSs, and in today's global world, this assumption does not apply to SLUs of English (Jenkins 2007; Pavlenko and Blackledge 2004). This point is particularly relevant to an EMI context where plurilingua-culturalism and diversity are common. Omitting the concept of SLU as a possibility for test-takers would most likely result in a needs analysis consisting of notions about test-taker needs that are off-target for these users. Such an approach can even suggest that those in power may principally be interested in molding test takers to fit their own language ideologies about a target situation (as discussed in 3.2.1 and suggested in 4.1.2). Taking these aspects of language use into consideration suggests that SLA views to language proficiency need to be adjusted to include alternative views that meet the needs of test takers whose primary aim may be to meet the demands of communicating efficiently in English in international contexts, where diversity and plurilingua-culturalism are prominent.

Using empirically based evidence, prominent scholars strive to conceptualize ELF as being an equally viable alternative to native-like English (e.g. Jenkins 2000, 2007; Mauranen 2006b, 2011a; Seidlhofer 2001, 2011). In taking a stance, ELF scholars argue against ELF being associated with the traditional SLA paradigm. Rather, they argue for concepts that are forming an emerging SLU paradigm. Taking this fundamental difference as the starting point, an examination of the SLA and ELF literature points to different proficiency targets for these two groups, as illustrated in Table 4-2. The left-hand column of Table 4-2 highlights key concepts associated with the proficiency targets of SLA and the right-hand column SLU.

**Table 4-2. What qualities measure proficiency of SLLs and SLUs?**

Traditional SLA paradigm	Emerging SLU paradigm
<b>Second language learner = SLL</b>	<b>Second language user = SLU</b> <i>Synonymous to ELF user</i>
<b>Idealized educated NS</b> <i>Correct, standard English for a homogenous speech community of monolingual NSs</i>	<b>Competent international speaker</b> <i>Appropriate English for a heterogeneous speech community of plurilingual-cultural NNSs (e.g. House 2003, Jenkins 2007)</i>
<b>Ability to imitate NS</b> <i>Aim to acculturate to L2 norms</i>	<b>Ability to accommodate NNS</b> <i>Aim to be understood and to maintain L1 identity (primarily through accent addition, Jenkins 2000)</i>
<b>Linguistic accuracy (main focus)</b> <i>Near NS grammatical, lexical, phonological competence</i>	<b>Resources of grammar &amp; expression</b> <i>Reasonably correct</i> Non-standard lexicogrammatical features: Regularized morphology (see Seidlhofer 2004) Overproductive morphology, words & phrases (i.e. youngness, maximalise; see Mauranen 2011b) Syntactic variants (Ranta 2006, 2009) <i>Plurilingual strategies</i> Code-switching, blending & mixing (i.e. drawing on shared languages) <i>Avoidance of idiomatic expressions</i> <i>Phraseological units</i> (i.e. linear expressions or chunks for managing organization and interaction, for distinction see Sinclair & Mauranen 2006)
<b>Communicative appropriateness</b> <i>Near NS competence</i>	<b>Communicative efficiency (main focus)</b> <i>Communication strategies</i> Repetition, metadiscourse, negotiating topic, enabling techniques (Mauranen 2005, 2006, 2007, 2009, 2011a) Strong 'other' orientation (Mauranen 2006, 2010, Kaur 2009) <i>Co-constructed utterances/interaction</i> Joint forces principle (Smit 2008) Let-it-pass principle (Firth 1990) Make-it-normal principle (Firth 1996) <i>Intelligibility</i> Lingua franca core (Jenkins 2000) Principle of clarity (Ahvenainen 2005) <i>Fluent interaction</i> The 'right' speed may not be the fastest Place emphasis on continuity and coherence (Hüttner 2009) <i>Smooth interaction</i> Backchanneling and latching (i.e. turns without pauses) Interactivity appears naturalistic (Hüttner 2009)
<b>Deviation view</b> <i>'Error' is a deviation from NS norms, caused by L1 transfer, interference, fossilization</i>	<b>Differential view</b> <i>'Difference' may signal an ELF variant that either enhances or hinders communication</i>
<b>Success determined by</b> <i>Conformity to centralized, standardized norms, i.e. imitative measures for language assessment</i>	<b>Success determined by</b> <i>Ability to accommodate/shift speech patterns to achieve effective communication that is context dependent</i>

A primary difference in these two proficiency paradigms lies in the view of the speaker being either a 'learner' or a 'user'. In opposition to the term 'learner', Firth argued for an alternative view of a lingua franca 'user', proposing a concept devoid of the traditional deficit model of NS communicative competence (Firth 1996: 241). Similarly, Jenkins

argues that the conceptual basis of ELF is not rooted in being a ‘foreign’ language to be attained by a ‘learner’ whose goal is to conform to NS norms (Jenkins 2006a: 139-140).

Another major difference between SLL and SLU is in the orientation towards a norm. In the SLA tradition, SLLs orient their linguistic and communicative targets towards a prestige variety of English, representing a homogeneous idealized ENL target culture. Against an idealized NS, their language use can be compared for accuracy and appropriateness. In contrast, ELF users speak English out of necessity and may be unfamiliar with each other’s cultural and linguistic backgrounds (Mauranen 2011a). Mauranen (2011b: 158) argues that NS conventions may not be available or appropriate in this context, and efficient communication may require the ability to deal with various cultures while performing a particular task. Moreover, speakers project their own identities, code-switch, utilize hybrid forms, and other similar acts to enhance efficient communication (e.g. Archibald et al. 2011; Mauranen 2012; Seidlhofer 2011; Söderlundh 2010). This unpredictable variation includes coping with different L1 accents and proficiency levels (e.g. Jenkins 2000, 2007). The variability aspect as well as having the ability to deal with it particularly illustrates how the criteria for adapting or ‘blending in’ are different for SLU and SLL. It also underscores different performance demands for SLL and SLU.

Because ELF contexts require plurilingua-cultural skills, ELF scholars advocate alternative norms for SLU. As an attainment goal, House (2003) asserts that the norm for SLU should be an ‘expert user’ instead of the NS. She defines this norm as “a stable multilingual speaker under comparable socio-cultural and historical conditions of language use, and with comparable goals for interaction”(ibid: 573). Alternatively, Jenkins (2009: 201) argues that the target of ELF speakers is the ability to accommodate each other by making adjustments to their English in a way that enhances mutual understanding. This view seems to suggest that the target for ELF users is not a particular norm – native or non-native speaker – but rather a mutual accommodation for enhanced communication. Similarly, Mauranen (2011b: 159) points out that a norm for communicative efficiency may be established by a group when no other linguistic authority prevails. An example of this type of norming is provided in Hynninen (2011; 2013). Her research illustrates speakers mediating norms that deliberately align with international practices rather than with British or American norms. This type of norming towards enhancing communication among ELF users would probably seem alien to monolingual NSs unaccustomed to such practices (ibid).

From the point of view of assessment criteria, the literature sharply contrasts on differences in the main communicative targets of SLL and SLU. For SLL, the main focus in spoken communication is native-likeness. Conversely, for SLU, the main focus is on communicative efficiency, where discourse features override grammatical accuracy (Mauranen 2011a). A prominent feature is communication strategies that indicate cooperativeness and facilitate comprehensibility (ibid) as well as constructing understanding (Mauranen 2006a). Examples of such strategies include fronting topics to enhance clarity (Mauranen 2011a), utilizing phraseological units to manage organization and interaction (Mauranen 2009a; Sinclair and Mauranen 2006), repeating information to increase lecture comprehension (Suviniitty 2012), and rephrasing to jointly construct



meaning (e.g. Cogo and Dewey 2006; Hüttner 2009; Smit 2010). The ELF literature suggests that the application of such communication strategies explains more about successful ELF communication than the accuracy of lexicogrammatical features (Mauranen 2011a). In addition, Hülmbauer (2009) suggests that ELF variants may form an important part of successful SLU communication. She views variants that are simple or that utilize features from shared languages as enhancers of communication more than as obstacles to it.

The SLA and SLU proficiency paradigms clearly differ in their views of what is important in terms of communicative success. This difference can also be seen in their treatment of what constitutes an ‘error’. ELF researchers commonly argue against a ‘deficit’ view for SLU of English and promote a ‘differential’ view instead (e.g. Firth 1996; Firth and Wagner 1997; House 2003; Jenkins 2007; Seidlhofer 2004). In the deficit view, SLL is measured against an idealized NS where deviations from the idealized homogenous ENL target culture indicate an ‘error’ (e.g. Corder 1981; James 1980; Selinker 1972). Alternatively, ELF scholars argue that ‘different’ should not be perceived as a sign of incompetence (e.g. Jenkins 2006a; Ranta 2009; Seidlhofer 2004). Rather, language that is ‘different’ may signal an ELF variant, a view that is supported by theories of language contact (e.g. Ehrenreich 2011; Mufwene 2001). Thus, variation in language use and communication is viewed as ‘different’ from ENL norms but not necessarily as constituting ‘error’. For example, Seidlhofer (2004) suggests that some lexicogrammatical elements of standard language ideology could be ignored in ELF communication, including what she terms ‘unilateral idiomaticity’, the third person –s ending, and a number of other features (see the list with examples in Seidlhofer 2004). She further argues for a target that is more achievable and relevant to the majority of ELF users, while pointing out that features of spoken ELF more closely resemble the English spoken by various L2 groups whose communications are said to outnumber ENL speakers by a large margin in global interchanges. In addition, Jenkins (2000) redefines phonological ‘error’, providing an alternative to deficit definitions common to an SLA paradigm. In 2007, she also advocates diverging from the NS as a linguistic point of reference for ELF (Jenkins 2007: 3).

Given these different proficiency paradigms, a rough checklist for determining test-taker needs could include the items in Figure 4-2. To establish the primary context of language use, a needs analysis, such as the one sketched here, could begin with the social basis to determine what aspects of test-taker performance might most usefully be assessed. Some key elements include whether the target context of use is a target culture or an international context, whether the target identity includes acculturating to NS norms or maintaining L1 identity, and whether efficient communicative targets include monolingual or plurilingual competence.

Determining test-taker needs: A rough checklist		
1.	What is the <i>target social context</i> , e.g. an ENL target culture or an international context of primarily ELF users?	
2.	What is the <i>target learner/user identity</i> , e.g. to acculturate to English norms or to maintain an L1 identity with the aim to be comprehensible in English?	
3.	What are the targets for <i>communicative efficiency</i> (below)?	
<i>Overall communicative effectiveness</i>		
	Near native speaker	Competent SLU
<i>Grammatical correctness</i>		
	Native-like	Reasonably correct
<i>Correctness of expression</i>		
	Native-like	Reasonably correct and easily understood, plus shows awareness of appropriate pragmatics
<i>Intelligibility</i>		
	Native-like	Non-native accent clear enough to be easily understood (i.e. represents a reasonable accent that could be expected by other users of the target community of practice), which includes an ability to adjust phonological features important for intelligibility as well as to comprehend different varieties of English
<i>Communication strategies</i>		
	Native-like	Effective use of accommodation strategies, including the ability to negotiate meaning and to employ multilingual approaches
<i>Fluency</i>		
	Native-like	Interactive fluency (i.e. ability to maintain unity and smooth interaction in discussion)

**Figure 4-2. A rough checklist for determining test-taker needs.**

Since performance testing must be based on situations relevant to test-taker needs, it is important to develop assessment criteria appropriate and relevant to the target audience of ‘learners’ or ‘users’. Without this criterion, there is the danger that test developers will suit their assessment criteria to their own notions about language proficiency. Given the differences between SLL and SLU of English, it is difficult to justify native-likeness as the basis for assessing ELF users. Even though some authorities may not agree with this view, it is important to develop assessment criteria that reflects the real-world needs of the parties involved rather than assess all L2 speakers with the same criteria, which are irrelevantly tuned and constructed for SLU.

### 4.3 Assessing language for specific purposes (LSP)

In order to appreciate the complexity of identifying language proficiency criteria related to EMI lecturing, it will be helpful to consider the professional needs and competencies valued in an international work environment. As the English-speaking community has become increasingly diverse linguistically, the needs for working life seem to have become even more blurred. In the academic domain, the multitude of ELF users suggests that the concept of identity has become more complex (e.g. Jenkins 2007), and that the notion of successful communication in the workplace has changed (e.g. Kankaanranta 2006; Kankaanranta and Louhiala-Salminen 2010; Kankaanranta 2012; Kankaanranta and Planken 2010). These ideas are also captured by Ehrenreich, who draws attention to changes in international business communication, where she argues that being socialized into this community of practice means “leaving behind traditional notions of appropriateness as experienced in EFL classrooms ... and adjusting their use of English to what is required and therefore ‘appropriate’ in these professional communities” (2009: 146). Her finding suggests that what is sociolinguistically appropriate is determined by the users, and that it has little to do with native-likeness.

In a complex multicultural environment, one of the arguments in this study is that the objective of EMI lecturers is to maintain their professional standards. According to Boswood (1999: 4-6), communicating to a professional standard, consists of mastery of non-verbal communication as well as mastery of the job-related knowledge and skills. On professional communication, the literature identifies intercultural competence as needed and desired in the workplace (e.g. de Graaff et al. 2007; Deardorff 2006; Huhta 1997; Huhta 1999; Niitamo 1999), including having a capacity for situation-sensitivity (e.g. McClelland 1973; Spensor and Spencer 1993), being able to manage ambiguity-diversity (e.g. Niitamo 2011) and to develop true cooperation (e.g. Nederström and Niitamo 2010), as well as being bicultural or bilingual (e.g. LaFromboise et al. 1993).

To create close ties with professional working-life needs, language teachers in higher education in Finland have actively interviewed prospective future employers and conducted projects with industry as a way to develop their LSP teaching. Having organized and supported this type of teaching for three decades, one could argue that Finland is among the leading contexts. At Finnish universities, Language Centres have been strategic in advancing LSP education through both training and active research (see, for example, Horppu and Lehtonen 2003; Karjalainen and Lehtonen 2005; Karlsson 1995; Rontu 2010; Sinkkonen 1998). Already in the 1970s, pioneering work in LSP was conducted by Ola Berggren, which significantly contributed to the development of needs assessment and oral assessment (1973, 1975, 1979). As part of this work, Berggren developed scales for assessing oral performance, which shifted the focus from general language skills anchored in native speaker competence in the mid 1970s to scales more occupationally/professionally oriented within a decade. At Finnish universities of applied sciences, active research has also advanced LSP teaching (see, for example, Airola 2004; Huhta 1999; Huhta 2010; Juurakko 2001; Takala 2001). One of their research projects, known as AMMKIA, produced validated assessment scales for LSP, and was one of the first applications of the CEFR in Finland (see Juurakko 2001: 13-14, 24, 35-36, 86). Given

these developments in LSP education in Finland, one could expect that new initiatives in higher education would be based on adequate knowledge of prior relevant work. It also raises the question of how informed and professionally adequate the in-house planning of the pilot-mentoring certification assessment was (see section 2.1).

Seeking to improve needs analysis tools, Huhta (2010) analyzed methods in seven language needs analyses used in industry and business to determine the fruitfulness of their returns to stakeholders. Her findings support using a variety of methods in empirically based language needs analyses. She also suggests a stronger focus on discourse and communication in professional communities of practice, including practices related to group interaction and intercultural communication. For LSP needs analysts, Belcher and Lukkarila (2011: 89-90) warn of the dangers of using preconceived needs-based analyses that do not tap into how SLLs see themselves and their future academic and professional needs. This danger also applies to assessing the communicative needs of ELF users, as indicated in section 4.2 as well as by the findings in Chapters 7 and 8.

Within the professional world, indigenous assessment practices form part of the professional culture, and can range from formal gatekeeping through examinations to informal built-in evaluation. These practices serve as a measure of professional competencies, which includes a specialized form of socialization. Through experience, individuals develop professional competence while simultaneously being socialized into specific contexts (Douglas 2000; Ehrenreich 2009; Jacoby 1998). Their experiences form the basis of a universal process involving social interaction, recognized as the medium through which culture-specific knowledge, skills, and practices are transferred and developed (e.g. Jacoby 1998; Lave and Wenger 1991; McNamara and Roever 2006; Ochs 1988; Schieffelin 1990; Schieffelin and Ochs 1986; Vygotsky 1978). Within an organization, the purpose of assessment is to follow professional development and to ensure quality control (e.g. McNamara and Roever 2006). This context suggests that determining the extent to which an individual has attained the required professional competence is a question for members of an indigenous group. In other words, to assess whether or to what extent an individual meets the professional standards of his or her specific context requires the knowledge of the indigenous members' inventory of assessment criteria, regardless of whether the individual is an L1 or L2 speaker (e.g. Garfinkel 1967; Geertz 1983; Goodwin and Goodwin 1996; Jacoby 1998; Jacoby and Gonzales 1991; Jacoby and McNamara 1999; Lynch 1985; Marlaire and Maynard 1993).

The importance of this indigenous aspect in assessing professional performance is acknowledged in the language proficiency literature (e.g. Douglas 2000; Douglas and Myers 2000; Jacoby and McNamara 1999; McNamara 1996). However, it has been difficult to operationalize using psychometric approaches. To date, the best framework for assessing LSP is the one proposed in Douglas (2000). In an attempt to find a working solution that avoids being either at the strong or weak end of the continuum in Figure 4-1, Douglas (2000) proposed a framework that allows for defining a psychological construct for assessing LSP. His framework builds on earlier work by J. B. Carroll (1968) and Clark (1972), from the area of work-sample assessment, and Bachman (1990) from the area of applied linguistics. From the viewpoint of assessment criteria, the framework incorporates 'indigenous assessment criteria' (cf. Jacoby 1998), defined as criteria used by 'subject

specialist informants' (cf. Selinker 1979) to assess the communicative performance of professionals in different fields. This approach increases the variability factor, which is not without problems and especially in highly specific contexts (for a full discussion see McNamara 1996; and Douglas 2000).

For LSP, indigenous assessment criteria can be derived from needs analyses based on any number of methods, including interviewing subject specialists, observing professionals at work, surveying organizations, and reviewing literature. A survey of the literature on EMI lecturing is limited but reveals the following. Suviniitty (2010; 2012) identified interactional features of lecture discourse, such as questions and repetition, to be a factor in making EMI lectures more comprehensible. In other words, EMI lecturers who used questions to elicit, check or seek information were easier to follow. In addition, EMI lecturers who emphasized points through lexical repetition also enhanced lecture comprehension (for discussion see, Suviniitty 2010, 2012). A study by Vinke (1995) on teaching behaviors in EMI also identified a number of features important to lecturing in this context, including redundancy, clarity and accuracy of expression, explicit structuring of information, and improvising skills. In her study redundancy referred to explaining ideas in different ways, including clarifying by restating, giving examples, and summarizing. Clarity and accuracy pertained to the expression of discipline-specific information as well as to the ability to paraphrase it. Improvising referred to the ability to address questions and comments from students spontaneously. These aspects of EMI lecture performance could be incorporated into indigenous assessment criteria for EMI lecturers, along with criteria collected from other suitable EMI research findings in LSP contexts.

From the viewpoint of developing assessment criteria for EMI professionals, the move towards incorporating indigenous assessment criteria in LSP assessments is positive. In this way, Douglas' model is a step forward over other psychometric models of language proficiency. However, like other psychometric models, its focus on measuring individual traits or qualities is disadvantageous for assessing ELF users where communicative efficiency relies on co-construction.

#### **4.4 Self-assessment in professional assessment**

Because self-assessment offers one approach to assessing performance-related competency skills, EMI lecturers in the present study were asked to self-assess their language competence for work-related purposes. A detailed review of the literature related to self-assessment is beyond the scope of this dissertation, thus this review aims to provide understanding of the role of self-assessment in professional assessment and of the need to support individuals in making realistic self-assessments.

In professional assessment, awareness of the need for self-assessment increased as economic theory and empirical research became interested in the individual and their social returns to education (Becker 1964; Psacharopoulos 1973; Schultz 1961). Since educational qualifications do not generally identify the various sets of non-cognitive skills an individual has acquired through formal education, the interest in such a measurement

arose with the aim to improve human capital in the labor market (Allen and van der Velden 2005). In the last few decades, the concept of human capital has become one of the driving forces behind economic development. This concept refers to the combined stock of competencies, knowledge, and social attributes embodied in an individual that form the ability to perform labor that produces economic value (ibid). This societal view directly connects education to professional competencies, as different social actors (from policy makers to students) have realized the importance of investing in training as a means to developing the stock of skills.

Although professionals are initially employed based on their educational and professional training, these qualifications mostly reflect skills from the cognitive domain. Thus, self-assessment provides a complementary means of assessing skills that are not addressed in traditional assessments well-suited to the cognitive domain (O'Malley and Pierce 1996). As a tool, self-assessment aids in skills measurement for performance competencies needed to fulfill particular tasks in a given situation as related to a particular job or position (Eva and Regehr 2005). Its function is to involve individuals in judging whether they have met an identified standard (ibid). For this reason, many educators and practitioners view self-assessment as an essential aspect of professional development (e.g. Graham 1988; Schön 1983; Zimmerman 1990). It develops competency through critical reflection of practice (e.g. Hine 2000), while promoting self-regulation (e.g. Laskey and Hetzel 2010; Zimmerman 1990), and encouraging self-directed, life-long learning skills (e.g. Levett-Jones 2005).

Despite its growth in popularity, the literature on self-assessment in health, education, and the workplace suggests its subjective estimates are inherently flawed in both laboratory and real-world settings (e.g. Ashton 2006; Dunning et al. 2005; Eva and Regehr 2005; Kane 1992). The correlation between views held and the behavior being measured was often moderate to poor, due to individuals claiming to have skills and attributes to a higher degree than what they possess (Dunning et al. 2005). This research suggests that acquiring an accurate view of oneself is an inherently difficult task for which individuals may not possess the critical information needed (ibid). To accurately self-assess, individuals need the ability to recognize their weaknesses and the advantages in admitting them (Brown 1990). The literature indicates that it cannot be assumed that individuals will be able to identify their shortcomings (Brown 1990, Dunning 2005), which signals the need for intervention.

On improving the accuracy of self-assessment, one common theme that manifested in the literature was utilizing information from other sources. For instance, it might be useful to provide feedback about an individual's strengths and weaknesses through alternative methods, such as peer review (Dunning, Heath, and Suls 2005). With intervention, an individual has a better chance of learning how to take steps towards drawing more realistic self-views about what they know about themselves (Dunning 2005). Other common themes for improving the accuracy of self-assessment include benchmarking in educational settings and having active boards of directors in the business world preventing CEOs from making costly mistakes based on faulty self-assessment (ibid).

Even though self-assessment has its drawbacks in measuring the characteristics of individuals, its popularity suggests that the potential value in the imperfect knowledge of skills and competencies is great (Allen and van der Velden 2005). It also suggests that the benefits of self-assessments outweigh the limitations. Because of the doubts about the validity and reliability of self-assessment methods, it is advisable to use more than one method of data collection.

## 4.5 Chapter summary

In this chapter, the literature review describes some of the complexities in developing assessment criteria that capture the real-world needs of test takers, with the social dimension being especially problematic. This point is partly expressed in McNamara's weak to strong performance hypothesis, where a strong performance predicts the ability to perform target tasks successfully and a weak one predicts the ability to use language in the target situation. This continuum makes a primary distinction where the former includes non-language factors while the latter does not. Although the most recent models of language assessment have moved away from the weakest type of performance by including some non-language factors, such as affective factors in the Bachman-Palmer (1996) model and specific purpose language ability in the Douglas (2000) model, these models still focus on measuring individual traits or qualities. This psychometric approach cannot account for co-constructed communicative efficiency, resulting in one-sided performances where the social context has been sidelined.

In addition, when viewing performance from a value-based perspective, the social dimension is subject to the ideals upheld by decision-making bodies, the acts of which could be seen as constructing language proficiency. To deter pre-conceived notions about test takers, it is important to conduct needs analyses into the current and future needs of test takers that include the concept of identity and heterogeneous target cultures, where plurilingua-culturalism may be a relevant facet. For ELF users, it would also be important to target their communicative needs for academic and professional purposes, which differ from the general linguistic competence of SLL. Establishing the contextual needs of ELF users would thus be a necessary step towards developing adequate criteria for assessing their performance.

On professional competence, previous studies emphasize the importance of measuring non-cognitive abilities that form part of the stock of competencies, knowledge, and social attributes necessary to perform a particular job well. They also recommend using a variety of empirically based methods to determine both the needs and the professional achievement or development of individuals. In particular, the literature suggests that self-assessment alone is insufficient for ensuring realistic assessments of individual performance. All in all, these factors point toward the social dimension as important to the development of assessment criteria for professional English in international contexts.

## **5 Reviewing the ideological basis of an assessment tool: CEFR**

The aim of Chapter 5 is to investigate the ideological basis of the CEFR from the perspective of the four widespread language ideologies described in Chapter 3. This critical review will reveal some of the challenges presented by language ideologies prominent in the 2001 version of the Framework for the assessment of professional spoken language in an international context. The investigation is from a theoretical perspective, and it serves to provide the basis for the analysis of the empirical data in Chapter 8. The reason for choosing the CEFR over other assessment tools is because five of its descriptor scales served as criteria for assessing the professional English of EMI lecturers in a pilot language-certification assessment at Aalto University in spring 2010. The research question under investigation examines the extent to which CEFR descriptors are a potential tool for assessing English for professional use in an international context.

This theoretical framework consists of five sections. First, an overview introduces the general background forming the basis of the CEFR. Following this, a discussion on the role of description in language is presented. The third section then examines six CEFR scales to discern what language ideologies are present in the descriptors, and the final section concludes the chapter.

### **5.1 CEFR – an overview**

Initiated by the Council of Europe, the CEFR supplies a common basis for language education throughout Europe. The purpose of this Framework was to create a reference work with a common core for language teaching, learning, and assessment. The idea was that this Framework would facilitate the comparison of language courses, curricula and proficiency in European countries. To achieve this goal, the Council of Europe established the CEFR, a document that describes “what language learners have to learn to do in order to use a language for communication and what [other] knowledge and skills they have to develop” (Europe 2001: 1). However, it does not dictate how language professionals should teach or what curricula should be imposed.

The CEFR is set up as a general framework of language reference that embodies a number of different aspects related to language education, not all of which were developed at the same time. As a tool developed over a ten-year period, the horizontal descriptive scheme and the vertical reference scales were not created as a unified concept. For example, work on the descriptor scales – a focal point in the present study – was conducted in the early 1990s. These descriptors thus depict approaches to language proficiency from even earlier times, which reflect an SLA paradigm (see Table 4-2). In contrast, the descriptive scheme was further elaborated in the late 1990s to include concepts promoted in European language policy, such as plurilingualism and pluriculturalism (see 2.3). However, no new descriptors or reference scales were created to



cover these (added) aspects. This asynchronous development results in a descriptive scheme that covers a broader range than the reference scales. Although the CEFR was not created initially as a single conceptual tool, it aims to provide a system that is as coherent as possible, while drawing on several sources. As development work continues, this framework could be viewed as a 'living' tool. How or whether its reference scales will later reflect concepts from an SLU paradigm is still unknown.

The CEFR is a theoretical framework constructed from a heuristic approach, allowing language users (testers, teachers, and learners) to refine it to meet their own needs. Its descriptive scheme has two dimensions: vertical and horizontal. The vertical dimension consists of six levels of communicative proficiency that are largely defined in terms of empirically derived estimates (i.e. consensus views) of difficulty based on either learners' or teachers' perceptions of language functions expressed as 'can do' statements. Such statements recognize the lower levels as having a role of functional importance in the language learning process (Hudson 2005: 215), and frames language learning and assessment in positive terms. The second dimension consists of a horizontal descriptive scheme dealing with L2 communicative competences and strategies, where the strategies "serve as a hinge between these competences (the learner's linguistic resources) and the communicative activities (what he or she can do with them)" (Little 2006: 168). Unlike the vertical dimension, the scales for the horizontal dimension are not the product of an independent empirical investigation (ibid: 168-69). Together, these two dimensions form the basis of the Framework, providing an approach for analyzing language in use in terms of strategies learners/users use to activate their *general* (i.e. non-linguistic) and *communicative language competences* in performing activities and learning *processes* that involve *productive* and *receptive* skills to construct discourse on particular *themes*, which enables them *to do tasks* under given *conditions* in *situations* that arise in various *domains* (Europe 2001: XV; emphasis in original). The italicized words represent the parameters for describing language use and ability to use language in the CEFR (ibid).

Its approach to language learning and assessment has strong merits, reflecting an integration of current concepts from SLA and consideration of guidelines for European language policy. The Council of Europe also encourages users to adapt this flexible framework to reflect local needs and contexts. To better cover various contexts, the Council has recently initiated a project to further develop the descriptors. However, the present critical review will illustrate the prominence of certain language ideologies in the descriptive discourse, which may limit the ability of the Framework to address the needs and the context of use for ELF users.

In the CEFR, successful performance at each reference level (i.e. A1 – C2) reflects an individual's competence. Competences are defined as "the sum of knowledge, skills and characteristics that allow a person to perform actions" (Europe 2001: 9). The CEFR distinguishes between communicative language competences, which include linguistic, sociolinguistic and pragmatic competences, and general competences, which cover non-linguistic factors such as knowledge (of the world), practical skills (e.g. social and professional), existential elements (e.g. affective factors), and ability to learn (for details, see Europe 2001: 101-108). Although the Framework presents no reference scales for

general competences, it contains a set for communicative language competences. Starting with the descriptive scheme, this latter competence is defined as follows:

- 1) Linguistic competences include lexical, phonological, syntactical knowledge and skills and other dimensions of language as system. (ibid: 13)
- 2) Sociolinguistic competence is concerned with the knowledge and skills required to deal with the social dimension of language use. (ibid: 118)
- 3) Pragmatic competences are concerned with the functional use of linguistic resources (production of language functions, speech acts), drawing on scenarios or scripts of interactional exchanges. (ibid: 13)

In particular, with regard to the pragmatic and sociolinguistic competences, what is stressed is “the major impact of interactions and cultural environments in which such abilities are constructed” (ibid: 13).

The CEFR has become widely accepted as the European standard for grading the language proficiency of an individual. To find out how well-known the CEFR is and how much it is used, the Language Policy Division of the European Commission conducted a survey in the spring of 2005 (Europe 2005). The results signal that the CEFR is “*rather widely known* in the responding institutions” (Europe 2005: 3, emphasis original), which included representatives from primary, secondary and higher education as well as examination providers (3.16 on a 0-4 scale), and it is “*quite widely used*” (2.24 on a 0-4 scale)” (ibid, emphasis original). The survey also reports that the common reference levels of language proficiency are the best known and most frequently used parts of the CEFR.

Given its asynchronous development and expansiveness, it could be expected that such a framework is likely to include some contradictions since it needs to be broad enough to be widely accepted, yet narrow enough to state specific aims. One point of contradiction lies in the discourse on the target language, which concentrates on NS-NNS contacts and seems to urge the learner to adapt to the target culture of the language being learned (Hynninen 2006). This, however, does not imply that the learner should aim to become native-like in the target/foreign language or achieve the unrealistic goal of NS competence. The CEFR states the following:

- 4) Level C2, whilst it has been termed ‘Mastery’, is not intended to imply native-speaker or near native-speaker competence. What is intended is to characterise the degree of precision, appropriateness and ease with the language which typifies the speech of those who have been highly successful learners. (Europe: 36)

Similarly, the CEFR claims “[l]evel C1 does not represent an ideal of unattainable perfection, but rather the highest level which it is practical to set as an objective for general language courses and public examinations” (ibid: 17). Although the CEFR makes these claims, the reference scales still tend to place NS competence as the highest form of language competence since level C2 skills are often compared to those of NSs:

- 5) Can hold his/her own in formal discussion [...] at no disadvantage to native speakers. (ibid: 78; level C2)

Moreover, only NS-NNS interaction is referred to in the CEFR scales in general:

- 6) Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. (ibid: 24, 27, 129; level B2)
- 7) Can sustain relationships with native speakers. (ibid: 35, 74; level B2)
- 8) Can understand any native speaker interlocutor. (ibid; level C2)

Thus, the CEFR appears to hinge on an NS-NNS axis, which implies that an ideal NS model is still in tact. It also points to the presence of NS language ideologies (a concept discussed in 3.5.3) as one guiding principle.

According to Hynninen (2006: 39-40), more contradictions center on plurilingual and pluricultural discourses<sup>14</sup>, which support diversity of language and cultures. Plurilingualism underscores the ability of learners to use different languages and to relate to other cultures. These accumulated competences form a common pool of resources that can be tapped into during intercultural encounters:

- 9) Plurilingual and pluricultural competence refers to the ability to use languages for the purposes of communication and to take part in intercultural interaction, where a person, viewed as a social agent has proficiency, of varying degrees, in several languages and experience of several cultures. (Europe: 168)

As Hynninen points out the pluricultural discourse above indicates that partial competences are also part of a learner's linguistic and cultural stock, which may include differing degrees of competence in different languages as well as varying degrees of cultural competence:

- 10) Plurilingual and pluricultural competence is generally uneven in one or more ways (Europe: 133)
- 11) Partial competence [...] is not a matter of being satisfied, for reasons of principle or pragmatism, with the development of a limited or compartmentalised mastery of a foreign language by a learner, but rather of seeing this proficiency, imperfect at a given moment, as forming part of a plurilingual competence which it enriches. (ibid: 135)

---

<sup>14</sup> *The CEFR defines the notion of plurilingua-culturalism as residing at the level of individuals, who do not store language "strictly in separate mental compartments". Instead, they build up their "communicative competence to which all knowledge and experience of language contributes and in which languages interrelate and interact" (Europe 2001: 4). This definition contrasts with multilingualism, which refers to different languages as being distinctly separate in a given society. In other words, they co-exist as separate entities and thus there is no switching, mixing, or blending (ibid).*

This type of discourse partly questions the NS model advocated in the descriptive scales and simultaneously promotes plurality in language teaching. Another way to view this would be to conclude that examples 9 and 10 include two separate discourses, one being NS-NNS and the other international discourse (Hynninen 2006: 41). In short, the reference level scales do not seem to reflect the full potential of the plurilingua-cultural discourse. Nor does the CEFR put forward a model that is unambiguous in its objectives. The next section briefly addresses the role of language description in the study of language use, bringing a contrastive view to the role of prescription (discussed in 3.5.2).

## 5.2 CEFR - Language description versus prescription

As discussed in section 3.5.2, attitudes to language may be expressed in prescriptive terms, where prescription entails value judgments related to correctness in language use. This view represents the position of public opinion (e.g. non-linguists). What has not yet been discussed is the role of description in language, which is typically the position taken by linguists who study language as a science. In this section, I briefly discuss description and prescription from a linguistic perspective, after which my aim is to examine the CEFR from these points of view in section 5.3.

From a linguistic perspective, language is studied through description, not prescription. All introductory textbooks on linguistics point out that linguistics is a descriptive science:

First, and most important, linguistics is descriptive, not prescriptive. A linguist is interested in what is said, not what he thinks ought to be said. He describes language in all its aspects, but does not prescribe rules of 'correctness'.

(Aitchison 1978: 13)

Consequently, the attitude of the linguist has little, if any, impact on the general public in terms of correct usage of language. Instead, authority is found in dictionaries, grammars, and similar reference books. The aim of linguists is to study language as objectively as possible. Modern linguists still agree that all forms of language are equal in principle.

Moreover, modern linguistics is based on a concept introduced by de Saussure (1915), the doctrine of arbitrariness. De Saussure realized that the linguistic forms that represent items in the real world do not necessarily have an inherent relationship to those referents. Thus, a term such as 'cat' in English is no better or worse way of referring to felines than equivalents in other languages. One can apply the same argument to grammatical systems. In short, if one wants to define a grammatical system with accurate descriptions, it would be difficult to form them based on value judgments of different languages and dialects. Hence, from the point of view of a linguist, it is not possible to show that one language or dialect is better than another based solely on linguistic grounds. In short, linguists do not participate in value judgments about language, whereas ordinary people (i.e. non-linguists) do.

Another concept perceived by De Saussure that is at the roots of modern linguistics is 'langue' (meaning language) and 'parole' (meaning speech). In de Saussure's dichotomy, langue approximately means language system and parole language use. A similar distinction was introduced in Chomsky (1965) as competence (the underlying rules of language innate in native speakers) and performance (actual language use). In both systems, *langue*, *competence* represents an abstract language system. In practice, much linguistic research has focused on the abstract element, giving consideration to universal ability to acquire a complex language system as opposed to researching the value judgments of particular usages considered important by public opinion.

At the level of language system, it would be difficult to argue that one language or dialect is linguistically superior to another. Nevertheless, a popular belief held by non-linguists is that official languages or standard varieties are superior to other forms (Collins 1999). Some also believe that linguistic superiority is the reason why one language spreads at the expense of another. At the level of language system, there is no way to demonstrate such a claim. Such claims are socially or politically motivated. Hence, one of the tasks of the sociolinguist is to explain why linguistic differences that are arbitrary are assigned social values. A challenging task for sociolinguists is to explain why people continue to use a non-standard dialect or variety when they acknowledge that the standard is the only correct form.

In light of these two views to language, let us consider the CEFR. By nature, is it descriptive or prescriptive? The CEFR appears to be primarily descriptive in nature with its scales being organized around language function rather than language form. Furthermore, the Framework acknowledges that language is in a continuous state of evolution as it is regularly used to meet the varying needs of its users for communicative purposes. It further acknowledges that no user of any complex and diversified language ever completely masters it (Europe 2001: 109). This view reflects the situation as it actually is in the real world. In spite of this, many nation states have set a standard language – a prescribed form that never reached the level of exhaustive detail (*ibid*) or even agreement among various authorities on what is 'correct' language (Milroy and Milroy 1999). Moreover, the linguistic description followed the same model employed for the corpus in use for the dead classical languages, a model abandoned by most linguists since they support describing language in use as it exists as opposed to what some authority thinks it should be (*ibid*).

However, none of the alternatively proposed models ever became generally accepted. Even though much work has been completed on linguistic universals, it has not produced results that directly benefit language learning, teaching, and assessment. The majority of descriptive linguists tend to practice codifying language, where they relate form and meaning using terminology that differs from the traditional models of description. This approach is also adopted in the CEFR:

It attempts to identify and classify the main components of linguistic competence defined as knowledge of, and ability of use, the formal resources from which well-formed, meaningful messages may be assembled and formulated.

(*ibid*: 109)

The scheme includes parameters and categories which may be useful in describing linguistic content and as a basis for reflection. Moreover, practitioners are free to use some other frame of reference if they wish. In that case, they need to specify the theory and practice which they are following. In this sense, the CEFR appears to free itself from prescriptivism. Although intended as a framework with the aim to provide a common frame of reference for an extensive range of contexts for language learning and assessment, closer examination of what the CEFR contains points toward prescriptivism (discussed in 3.5.2) and native-speakerism (discussed in 3.5.3) as the primary guiding principles upon which its ideological base resides.

With prescriptivism and native-speakerism at its roots, the CEFR can be viewed as a hierarchical scheme that treats language learning and assessment in decontextualized terms, where language is measured in relation to a set of predetermined and fixed norms. Close inspection of the Framework appears to imply that language assessment is largely depicted by a culture of 'correctness' as defined by NS norms, as illustrated in sections 5.3.2 and 5.3.3. The concerns here are with 'NS normativity' as related to NS language ideologies and with the prescriptive attitudes to language 'correctness', which largely pertain to standard language ideologies where prescriptive judgments about 'correct' and 'incorrect' use of language are imposed by those above, referred to as the language guardians or gatekeepers (cf. Milroy and Milroy 1999; Jenkins 2007). Such views relate to language standardization, which is based on the acceptability rules of NS written language. In this respect, a number of descriptors in assessment scales, such as those in the CEFR, may be counter-productive for assessing English as an international language (Seidlhofer 2003).

### **5.3 Examining the CEFR for language ideologies**

The CEFR claims to be comprehensive: "it should attempt to specify as full a range of language knowledge, skills and use as possible . . . and that all users should be able to describe their objectives, etc., by reference to it" (Europe, 2001: 7). However, many of the concepts introduced are not actually incorporated into the scales. In fact, users are encouraged to provide additional information, specific to their own situations, as indicated in the advice boxes:

Users of the Framework may wish to consider and where appropriate state . . .  
(ibid: 2001)

The advice box appears throughout the text, of which the following is relevant to the current discussion:

Communicative language competences including linguistic, pragmatic and sociolinguistic competences.  
(ibid: 108-130)

In this way, the CEFR appears to be organized in a manner that is less arbitrary than other proficiency descriptor scales available.

Yet, a closer examination of this heuristic approach reveals discourse that relates its reference levels to language ideologies that contrast with plurilingua-culturalism. For instance, the CEFR describes common reference levels for qualitative aspects of spoken language with descriptors under headings, such as accuracy and control, both of which signal prescriptivism. It also introduces linguistic normativity of communication by taking the NS as its standard measure, a practice common in ELT and based on NS language ideologies. This inherent normativity employs a presupposition of communication in NS-NNS interaction where the aim is to acculturate to NS norms. This view implies beliefs related to monolingualizing language ideologies, including one-nation-one-language ideologies. It simultaneously contrasts with NNS-NNS interaction where plurilingua-culturalism defines the international context of usage. The following sections provide illustrative examples that support this point. In this section, I will examine some of the communicative language competences, starting with linguistic competences and then moving to sociolinguistic and pragmatic competences.

### **5.3.1 Linguistic competence**

In the CEFR (ibid: 108-109), the discourse describing linguistic competence follows recent descriptive linguistic practices. It introduces linguistic competence as a formal system of language through which meaning is expressed. It also acknowledges that no complete description of any language exists and that no language is never completely mastered. Because descriptive linguists have not found the concept of linguistic universals to be useful for ELT, the approach adopted in the CEFR is an attempt to codify practice, “relating form and meaning” (ibid: 109). To achieve this, the CEFR uses traditional terminology except in cases that fall outside that range. It attempts

[To] identify and classify the main components of linguistic competence defined as knowledge of, and ability to use, the formal resources from which well-formed, meaningful messages may be assembled and formulated.

(ibid: 109)

The CEFR distinguishes six types of competences and their scales: grammatical competence, lexical competence, phonological competence, semantic competence, orthographic competence, and orthoepic competence. Of these, only the first four are relevant to the current discussion.

### 5.3.1.1 On grammatical competence

In the CEFR, grammatical competence is defined as "knowledge of, and ability to use, the grammatical resources of a language" (Europe, 2001: 112). This competence entails the ability to understand and to assemble meaningful, well-formed phrases and sentences as opposed to memorized sentences reproduced as fixed formulae.

Table 5-1 shows the CEFR descriptors for the common reference levels provided under the heading 'grammatical accuracy'.

**Table 5-1. The CEFR scale of grammatical accuracy. Source: Council of Europe 2001: 114, emphasis mine, where 1) bold, red text denotes words related to prescriptivism, 2) bold, black text signals phrases related to communication, and 3) underlined segments imply native-likeness.**

Grammatical Accuracy	
C2	<u>Maintains consistent grammatical <b>control</b> of complex language, <b>even</b> while attention is otherwise engaged (e.g. in forward planning, in monitoring others' reactions).</u>
C1	<u>Consistently maintains a high degree of grammatical <b>accuracy</b>; <b>errors</b> are rare and difficult to spot.</u>
B2	Good grammatical <b>control</b> ; occasional 'slips' or non-systematic <b>errors</b> and minor <b>flaws</b> in sentence structure may still occur but they are rare and can often be <b>corrected</b> in retrospect. <u>Shows a relatively high degree of grammatical <b>control</b>. <b>Does not make errors which cause misunderstanding and can correct most of his/her mistakes.</b></u>
B1	Communicates with reasonable <b>accuracy</b> in familiar contexts; generally good <b>control</b> though with noticeable <b>mother tongue influence</b> . <b>Errors</b> occur, <b>but it is clear what he/she is trying to express</b> . <u>Uses reasonably <b>accurately</b> a repertoire of frequently used 'routines' and patterns associated with more predictable situations.</u>
A2	Uses some simple structures <b>correctly</b> , but still systematically makes basic <b>mistakes</b> - for example tends to mix up tenses and <b>forget to mark agreement</b> ; nevertheless, <b>it is usually clear what he/she is trying to say</b> .
A1	Shows only limited <b>control</b> of a few simple grammatical structures and sentence patterns in a memorized repertoire.

In this scale for grammatical accuracy, there are several terms that point to prescriptivism as the measure of successful communication, such as 'accuracy', 'correct', 'control', 'mistake', 'error', and 'mother tongue influence'. Moreover, these terms appear frequently throughout the Framework and prominently in the scales, a point previously raised in the literature (Dewey 2009). In the A2 descriptor in Table 5-1, the term 'correct' clearly points towards correct usage of the language, which we can assume would be judged according to a codified handbook of grammar. Furthermore, the term 'accuracy' in the C1 and B1 descriptors can be understood as a synonym for 'correct'. In addition, the term 'grammatical control' seems to be another way to express 'accuracy'. The remaining three terms, 'mistake', 'error', and 'mother tongue influence' appear to be antonyms suggesting lack of correctness, with 'mother tongue influence' serving as a clear marker. From the perspective of language assessment, we can assume that these words and phrases will be interpreted to mean that the grammaticality of a speaker's use of English is measured according to the norms of standard language. Not only do terms like 'accurate', 'correct', and 'control' carry prescriptive tones, but also the conceptualization of the term 'error' in the CEFR:



Errors are due to an '*interlanguage*', a simplified or distorted representation of the target competence. (ibid: 155, emphasis in original)

In addition, phrases at levels C1-C2, such as “maintains consistent grammatical control of complex language” (ibid: 114), imply native-like competence. Although the term ‘native speaker’ is not explicitly stated in the CEFR scale for grammatical accuracy, the following statement does accompany the scale:

The syntax of the language of a mature native speaker is highly complex and largely unconscious. (Europe, 2001: 115)

Not only is the native speaker a vague concept, there are other problems with trying to use it as the measure for ELF. The term 'native speaker' belongs to a paradigm that views NNSs as language ‘learners’, not as language ‘users’ (as discussed in 4.2). Thus, according to SLA research, a SLL progresses through developmental stages along the path to native-like language. This view of L2 ‘learners’ is explicit in Selinker's interlanguage theory (1972; 1992), where the idealized speaker is a NS and the ultimate goal of the learner is to attain NS proficiency. To fall short of that mark results in either *interlanguage* (in the event that learning is still ongoing) or *fossilization* (in the event that learning has ended), despite the fact that estimates of native-likeness range from one percent (Bley-Vroman 1989) to five percent (Selinker 1972). In spite of this, the CEFR defines the cause of errors as 'interlanguage'. Furthermore, no consideration is given to whether English is spoken as a lingua franca in international contexts or used in ENL monolingual settings, even though ELF, unlike ESL varieties, is a contact language across language groups (see discussion in 2.4). It seems to me that a better definition for ‘error’ involving NNS-NNS contact would be ‘a word or sequence that does not have support from the linguistic communities in which the speaker wishes to mix’. This definition concerns itself with natural language use and is free of the deficit views that stem from NS language ideologies present in SLA definitions of ‘error’ and ‘mistake’, based on Chomsky’s idealized educated NS.

### 5.3.1.2 On lexical competence

The CEFR first presents a definition of lexical competence that meets its general objective of being open and non-dogmatic: the definition cannot be tied to any specific language ideologies or educational theories. It defines lexical competence as “knowledge of, and ability to use, the vocabulary of a language” (Europe 2001: 110) and it consists of two components: lexical and grammatical. These two components have separate descriptors for the common reference levels, of which the former is vocabulary range (see Table 5-2) and the latter vocabulary control (see Table 5-3).

For grammatical elements, the CEFR presents a list of closed word classes (ibid), all of which are representative of the structure of English. Some of them, however, may not be relevant to ELF usage. For example, ELF research has shown the following

lexicogrammatical structures to be generally unproblematic for communicative success in NNS-NNS interaction:

- ‘Dropping’ the third person present tense –s,
- ‘Confusing’ the relative pronouns who and which,
- ‘Omitting’ the definite and indefinite articles where they are obligatory in native speaker language use

(Seidlhofer 2003: 18)

Likewise, the lexical elements, which include fixed expressions consisting of groups of words learned and used as a whole, tend to be prevalent in NS usage but less common in NNS-NNS interaction. The fixed expressions include phrasal idioms and verbs, fixed collocations, proverbs, and other similar expressions that are characteristic of metaphorical language use, such as ‘he drove hell for leather’ or ‘he kicked the bucket’. Such idiomatic expressions are culture specific and signal the presence of NS language ideologies. While such descriptors may be relevant to SLL, the ELF literature does not support this view. As Seidlhofer (2004) explains, it may be that the social behavior representative of NS models and identities are not operable in NNS-NNS interaction and that certain NS norms may be viewed as suspenseful. This difference suggests that descriptors for lexicogrammatical appropriateness must be based on alternative criteria, such as those representative of an SLU paradigm (see discussion in 4.2).

In SLA research, the use of formulaic speech (termed ‘fixed expressions’ in the CEFR) has been a recurring topic. It is a known fact that a language consists of various kinds of formulaic speech that range from simple patterns to more complex ideas. SLA research has attempted to uncover the patterning and regularity of such formulaic speech. Like current linguistic models, the CEFR sees formulaic speech as either primarily lexical or grammatical bundles. These ‘fixed expressions’ contain fixed and variable parts, and tend to be fairly easy to recognize but not so easy to define (Mauranen 2009a). While they have been discernible in the speech of NSs, they are more difficult to spot in the speech of NNSs (ibid). Moreover, previous research has noted this difference as a marker of attainment: L2 speakers at advanced levels of language competence are differentiated from NSs by their inability to use such ‘fixed expressions’ in a native-like manner (Mauranen 2009a; Nattinger and Decarrico 1992; Pawley and Syder 1983; Wray 2002). One problem with this approach is that it caters to NS language ideologies. It also carries prescriptive tones that presuppose canonical forms on learners and users although recent research has shown that even NSs do not restrict themselves to the preferred forms (Mauranen 2003). A third problem is the lack of relevance for measuring successful communication among ELF users (see discussion in 4.2).

Against this background, I will now examine the lexical scales. Starting with the scale for vocabulary range, shown in Table 5-2, the descriptors from levels B2 to C2 contain key terms of interest to the current discussion.

**Table 5-2. The CEFR scale of vocabulary range. Source: Council of Europe 2001: 112, emphasis mine, where bold, red text signals one of the following: prescriptivism, native-likeness, or everyday contexts (as opposed to LSP).**

Vocabulary Range	
C2	Has a good command of a very broad lexical repertoire including <b>idiomatic expressions</b> and <b>colloquialisms</b> ; shows awareness of connotative levels of meaning.
C1	Has a good command of a broad lexical repertoire allowing <b>gaps</b> to be readily overcome with circumlocutions; little obvious searching for expressions or <b>avoidance strategies</b> . Good command of <b>idiomatic expressions</b> and <b>colloquialisms</b> .
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 <b>lexical gaps</b> can still cause <b>hesitation and circumlocution</b> .
B1	<b>Has a</b> sufficient vocabulary to express him/herself with some <b>circumlocutions</b> on most topics pertinent to his/her <b>everyday life</b> such as family, hobbies and interests, work, travel, and current events.
A2	Has a 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.

This scale includes several terms linked to NS language ideologies that would impact how language assessment would be evaluated. First of all, at levels C1 and C2, the terms 'idiomatic expression' and 'colloquialism' are indicative of NS language ideologies as central to the measurement of success. In addition, at the B2 level, the terms 'lexical gaps' and 'hesitation' are keys to assessing the level. The problem here is that such terms carry negative tones although phenomena such as 'hesitations' and 'dysfluencies' are even common in NS spoken language (see, for example, Biber et al. 1999: 1066-68). This approach implies spoken language is described in terms of deviations from written language, a practice that sets standard language ideologies as the guiding principle for measuring lexical competence.

For the assessment of vocabulary control, the CEFR presents the descriptors in Table 5-3. Here we find many of the same terms as in Table 5-1 (on grammatical accuracy). Common to the whole scale, except level A1 which has no descriptor, are the words: 'correct', 'incorrect', 'error', 'accuracy' and 'control'. Similar to the criteria for grammatical accuracy, these terms signal the presence of standard language ideologies as the measure of competence.

**Table 5-3. The CEFR scale of vocabulary control. Source: Council of Europe 2001: 112, emphasis mine, where bold, red text denotes words related to prescriptivism.**

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

For ELT professionals, the general attitude to 'errors' in language learning and assessment are captured in the title of this ELT resource book, *Learner English: A Teacher's Guide to Interference and Other Problems* (Swan et al. 2001), which maintains a prominent position on reading lists for in-service teacher training programs (Dewey 2009). Such resource books have at their roots the prescriptive voice of grammarians like Randolph Quirk (1985). As Jenkins notes:

There is really no justification for doggedly persisting in referring to an item as 'an error' if the vast majority of the world's L2 English speakers produce and understand it.

(Jenkins 2000: 160, emphasis original)

Not only is the CEFR view to 'error' problematic, but also the question of how to decide what constitutes an error and what equates language innovation. In SLA research, an aberration found in the lexical items produced by a NNS tends to be labeled as 'overproductive', 'random error', or 'idiosyncratic' (Corder 1981; Selinker 1992). This view labels any L2 difference as L1 interference, even though contrastive analysis (CA) predicts errors that have never been observed in L2 utterances. CA is also incapable of explaining the uniformity of errors made by a variety of L2 speakers. Rather than follow these NS ideologically based concepts for what constitutes a 'lexical error', a more naturalistic view of SLU needs consideration. For instance, when a NS uses a new word, it may be 'coined' and put into use. In other words, it is viewed as linguistic innovation. A similar phenomenon also occurs in NNS speech, especially when a non-standard lexical item is coined to fill a lexical gap, for example, for items that have no English word. This is also linguistic innovation. Throughout the history of language, words have been invented, partly due to contact with other languages and partly due to the need for new words, among other reasons (see, for example, Bradley 1904). Language change occurs through speakers who innovate. Thus, speaker innovations are a driving force in linguistic variation. As this example illustrates, the question of what constitutes a lexical 'error' needs serious re-consideration, especially for evaluating the SLU of English, where the

prescriptive rules of standard language ideologies seem archaic and unjustifiable. Similarly, prominent scholars have argued this point, of which the most notable include Jenkins (e.g. 2000, 2007), Mauraen (e.g. 2012), and Seidlhofer (e.g. 2011).

### 5.3.1.3 On phonological competence

In describing phonological competence, the CEFR discourse clearly signals NS language ideologies as the guiding principle. The initial description, however, seems rather neutral: it includes encompassing knowledge of and skill in perceiving and producing the sound units (e.g. phonemes and allophones) of a language, the distinctive features of phonemes, the syllable structure and composition of words including word stress and word tones, the prosody of sentence phonetics (e.g. sentence stress, rhythm, and intonation), as well as phonetic reduction (e.g. vowel reduction, strong and weak forms, assimilation, and elision) (Europe, 2001: 116).

What is notable is that all of the phonological items in the above list are features that are taught in models where the NS is the point of reference for correctness. Such a target seems disadvantageous to the communicative aims and needs of ELF users, as shown in the work of Jenkins (2000). This point is also raised in Kirkpatrick (2007), whose research shows syllable-timed, local varieties of English to be more easily intelligible than the stress-timed ENL varieties (ibid 2007, see Chapters 9 -11). Other research also argues against the NS as the target measure, deeming it as unattainable and unrealistic (e.g. Cook 1999), and especially in outer circle countries (e.g. Kachru 1985), making any argument about its relative international intelligibility as irrelevant (e.g. Jenkins 2000, 2007; Kirkpatrick 2007).

As an alternative to the NS model of phonology, Jenkins (2000) proposes a lingua franca core (LFC) as a pronunciation syllabus for international English, where SLU dominates. Her seminal work renders some phonological items typical to assessment scales with native-like targets, such as those in the CEFR, as unimportant for successful communication. As crucial to intelligibility, Jenkins outlines three main areas: segmentals, suprasegmentals, and articulatory settings. At the segmental level, that is, sound units, she argues that the ability to produce most of the consonants of English is important for SLU. Only two consonants are omitted from the lingua franca core, the dental fricatives /θ/ (theta), and the velarized /l/. In final consonant clusters, Jenkins argues that elision is permissible for ease of articulation even though pronunciation books, such as *Elements of Pronunciation* (Mortimer 1985), insist that learners produce full consonant clusters. On phonetic reduction, Jenkins points out that weak forms are seldom learned although often taught in EFL classrooms in Britain (and here I could add the USA). On suprasegmentals, she states that weakening one form does not necessarily highlight another element. Among the suprasegmental features, Jenkins argues that nuclear stress is crucial for intelligibility. On articulatory settings, she considers the holistic factors in producing speech where she argues that aspects of articulatory settings and voice quality are very important phonological elements of intelligibility for English in NNS-NNS contact. Note that this latter point, voice quality, is missing from the CEFR. From the perspective of language

ideologies, what is notable about Jenkins’ LFC is that it does not pit intelligibility against native-likeness. Rather, it targets a comprehensibility goal that utilizes plurilingualism as a basic notion for assessing the phonological control of spoken English for international communication.

In addition to the list of phonological elements in the beginning of this section, the CEFR also provides a scale of descriptors for the common reference levels under the heading 'phonological control', as shown in Table 5-4.

**Table 5-4. The CEFR scale of phonological control. Source: Council of Europe 2001: 117, emphasis mine where bold, red text denotes words that point to NS language ideologies or prescriptivism.**

Phonological Control	
C2	As C1
C1	Can vary intonation and place stress <b>correctly</b> in order to express finer shades of meaning.
B2	Has acquired a clear, <b>natural</b> , pronunciation and intonation.
B1	Pronunciation is clearly intelligible even if a <b>foreign</b> accent is sometimes evident and occasional <b>mispronunciations</b> occur.
A2	Pronunciation is generally clear enough to be understood despite a noticeable <b>foreign</b> accent, but conversational partners will need to ask for repetition from time to time.
A1	Pronunciation of a very limited repertoire of learnt words and phrases can be understood with some effort by <b>native speakers</b> used to dealing with speakers of his/her language group.

As shown in the table, this scale contains several terms related to NS language ideologies, including ‘native speaker’, ‘foreign’, ‘mispronunciation’, and ‘correctly’. The choice of the term ‘foreign accent’ in the B1 and A2 levels rather than ‘accent’ is indicative of a NS model as the measure of control and accuracy. All speakers of a (spoken) language have an accent of some kind whether local, standard, or foreign. In the A1 descriptor, the term ‘native speaker’ contaminates the whole scale for any assessment involving NNS-NNS interaction. In an NS-NNS context, an assessor would most likely approach assessment from the viewpoint of a prestige accent. Since the CEFR is intended to support intra-European linguistic and cultural diversity, we can assume that the NS model for English is British English. In England, the primary teaching model for SLLs and the one against which they should be judged is Received Pronunciation (RP), an accent spoken by less than three percent of the British population (Jenkins 2000: 14). In other words, a very small group is setting the standard (ibid). Moreover, someone who speaks English effectively, but whose pronunciation is considered ‘wrong’ may find his/her social mobility blocked. Attitudes to pronunciation (prescriptive in nature) have been known to lead to discrimination on linguistic grounds, which is publicly acceptable (Milroy and Milroy 1999: 2). While these examples illustrate the pressure for SLLs to conform to NS norms as well as for all speakers to conform to standard language, they do not exemplify the comprehensibility goals of ELF users, where plurilingualism can be an asset. This difference points toward contrasting language ideologies as a basis for assessment: NS verses English-as-a-global-language ideologies.

### 5.3.2 Sociolinguistic competence

The second communicative competence in the CEFR is sociolinguistic competence. Although the general definition of this concept omits discourse that would link it to any of the widespread language ideologies in Chapter 3, the same cannot be said about its descriptor scale. I will first discuss the general definition, after which I will examine the scale of sociolinguistic appropriateness for discourse related to specific language ideologies.

The CEFR definition for sociolinguistic competence is quite neutral in that it is devoid of discourse related to specific language ideologies:

Sociolinguistic competence is concerned with the knowledge and skills required to deal with the social dimension of language use.

(Europe 2001: 118)

The CEFR also mentions sociocultural phenomena as being relevant to this competence. The matters treated under the CEFR heading 'sociolinguistic competence' pertain to language use that is not dealt with elsewhere, such as politeness conventions, register differences, linguistic markers and so on.

Table 5-5 shows the CEFR descriptors for the common reference levels provided under the heading 'sociolinguistic appropriateness' (Europe 2001: 122). Thus, the CEFR provides a classificatory tool for sociolinguistic competence with parameters contingent on 'appropriateness', which is in line with traditional methods of linguistic description. It is common knowledge that language differs in being appropriate for different purposes and different situations.

Although the concept of appropriateness has been widely used in language education in discussions on sociolinguistic variation, it has been at the center of controversial policy on language education in Britain in recent years. Moreover, Fairclough (1995: 233) argues that theories of appropriateness reinforce controversial educational policies on language teaching, including a "competence-based communication skills view" where the emphasis is on spoken language.

Fairclough (ibid: 234) sees appropriateness as an ideological category "linked to particular positions within a politics of language – within a struggle between social groups in a speech community for control of (or 'hegemony' over) its sociolinguistic order". He exemplifies his point with the following example from the Cox Report (1989):

Pupils working towards level 7 should consider the notion of appropriateness to situation, topic, purpose and language mode and the fact that inappropriate language use can be a source of humour (either intentional or unintentional) or may give the impression that the speaker or writer is pompous or inept or impertinent or rude. Pupils should learn that Standard English is the language of wide social communication and is particularly likely to be required in public, formal settings. Teaching should cover discussion of the situations in which and purposes for which people might choose to use non-standard varieties rather than Standard English, e.g. in speech with friends, in

a local team or group, in television advertising, folk songs, poetry, dialogue in novels or play.

(6.29, original italics; quoted in Fairclough 1995: 234)

With this example, Fairclough points out that appropriateness is the foundation upon which the Cox Report builds its policy for the teaching of Standard English. Moreover, the prescriptive nature of the term appropriate becomes clear in discussions on inappropriateness. In the above extract, inappropriateness is depicted as a source of humor as well as behavior leading to unfavorable social judgments, as indicated by the words 'inept' and 'rude'. However, no serious reasons for inappropriateness, such as racist comments or humor, are presented.

Fairclough further points out that the Cox Report ties appropriateness to standard language ideologies in the following statement:

We need both accurate descriptions of language that are related to situation, purpose and mode (i.e. whether the language is spoken or written), and prescriptions that take account of context, appropriateness and the expression of meaning.

(4.19; quoted in Fairclough 1995: 236)

In this way, the Cox Report links description with prescription, which separates appropriateness from such concepts as 'correctness': appropriateness is prescribed in accordance with descriptively established practices in the speech community (Fairclough 1995). A problem with this approach is that it assumes that speech communities exist in a static state. However, sociolinguistic order is not so clear-cut. Linguistic variation lives in a dynamic state where the relation between context and purpose may be indeterminate. Models of language variation based on appropriateness presuppose clear-cut conventions, which project a misguided and unsustainable image of how sociolinguistic orders are structured (ibid). The report also goes a step further: it claims that it is possible to expand the repertoires of pupils by adding Standard English to their existing dialects, and it rests its case on the concept of appropriateness (4.43 Cox Report 1989). Closer examination of the report reveals that only Standard English is appropriate in public, formal domains, in other words, those domains that have the most prestige. However, it is questionable whether this can be achieved without erasing existing dialects or languages.

The above examples show how appropriateness has been central to the rethinking of language education in Britain, which embraces competence-based views for teaching communication skills. They illustrate the power of authority and treat hierarchy as natural. They also show how appropriateness can be used for political and ideological purposes. More specifically, they demonstrate how appropriateness may be used as a vehicle for policies on the teaching of Standard English and to dismiss shared language values of a particular dialect.

Against this background, I will now examine the scale for sociolinguistic appropriateness in Table 5-5. The lower levels of the scale, A1 and A2, make no overt mention of appropriateness. However, the mention of NS at the B2 level implies that English embodies a specific national culture; presumably British – even though English



embodies a number of different national cultures. At the A1-A2 levels, it seems that the assumption is that L2 speakers have grasped the basic sociolinguistic skills for sociocultural encounters in NS-NNS interaction. By level B1, L2 speakers need to demonstrate sociolinguistic competence that is appropriate, especially where politeness conventions are concerned. Progressing up the scale to level B2, the descriptors begin to contain some questionable language, closely linked to NS normativity.

**Table 5-5. The CEFR scale of sociolinguistic appropriateness. Source: Council of Europe 2001: 122, emphasis mine, where bold, red text denotes words related to prescriptivism or NS language ideologies.**

Sociolinguistic Appropriateness	
C2	Has a good command of <b>idiomatic expressions</b> and <b>colloquialisms</b> with awareness of connotative levels of meaning. Appreciates fully the sociolinguistic and sociocultural implications of language used by <b>native speakers</b> and can react accordingly. Can mediate effectively between speakers of the target language and that of his/her community of origin taking account of sociocultural and sociolinguistic differences.
C1	Can recognize a wide range of <b>idiomatic expressions</b> and <b>colloquialisms</b> , appreciating register shifts; may, however, need to confirm occasional details, especially if the accent is unfamiliar. Can follow films employing a considerable degree of <b>slang</b> and <b>idiomatic usage</b> . Can use language flexibly and effectively for social purposes, including emotional, allusive and joking usage.
B2	Can with some effort keep up with and contribute to group discussions even when speech is fast and <b>colloquial</b> . Can <b>sustain relationships with native speakers without unintentionally amusing or irritating them or requiring them to behave other than they would with a native speaker</b> . Can express him or herself <b>appropriately</b> in situations and avoid <b>crass errors</b> of formulation.
B1	Can perform and respond to a wide range of language functions, using their most common exponents in a neutral register. Is aware of the salient politeness conventions and acts <b>appropriately</b> . Is aware of, and looks out for signs of, the most significant differences between the customs, usages, attitudes, values and beliefs prevalent in the community concerned and those of his or her own.
A2	Can perform and respond to basic language functions, such as information exchange and requests and express opinions and attitudes in a simple way. Can socialize simply but effectively using the simplest common expressions and following basic routines. Can handle very short social exchanges, using everyday polite forms of greeting and address. Can make and respond to invitations, suggestions, apologies, etc.
A1	Can establish basic social contact by using the simplest everyday polite forms of: greetings and farewells; introductions; saying please, thank you, sorry, etc.

At level B2, the term 'crass errors' appears in the descriptor, which would most likely influence how assessment would be approached. Following SLA, the tradition for the assessment of 'error' resides in NS language ideologies, where the goal of L2 learning is largely defined by the avoidance of deviation. Within this ideology, the SLL tends to be cast in a negative light as emphasized by the following B2 descriptor:

Can sustain relationships with native speakers without unintentionally amusing or irritating them or requiring them to behave other than they would with a native speaker.

(ibid)

Not only does it cast a negative light on the L2 speaker, it also carries ideological tones of appropriateness similar to those expressed in the Cox Report. It is clear that inappropriate language may become a source of irritation or humor for the NS. While the Cox report upholds standard language ideologies as the measure of appropriate behavior for all speakers of English (L1 and L2 alike), the CEFR reinforces it as a measure for L2 speakers through NS language ideologies. This practice applies at levels B2 and higher, where L2 speakers must demonstrate an ability to make jokes viewed as appropriate with NSs. At the upper CEFR levels, the key words for demonstrating sociolinguistic appropriateness include 'idiomatic expressions' and 'colloquialisms', which reinforce culture-specific NS norms. Such definitions of sociolinguistic appropriateness are unlikely to benefit ELF users, who seldom interact with NSs and who may be unfamiliar with ENL cultures. Nor do such definitions capture the essence of appropriateness in terms of plurilingua-culturalism, an ideology present in the CEFR descriptive scheme but missing in its reference scales.

### 5.3.3 Pragmatic competence

The third communicative language competence in the CEFR is pragmatic competence. As with the previous two competences, the CEFR describes this competence in broad, general terms. The description includes three types of competences, all of which are concerned with the user/learner's knowledge of how messages are:

- a) Organized, structured and arranged (i.e. discourse competence);
- b) Used to perform communicative functions (i.e. functional competence);
- c) Sequenced according to interactional and transactional schemata (i.e. design competence)

(ibid: 123)

For these three competences, a variety of scales are available under the following headings: flexibility, turntaking, thematic development, coherence and cohesion, propositional precision, and spoken fluency. Of these, two scales are included in the present discussion since they served as measurement tools for EMI lecturers (in the performance-based pilot-certification assessment): spoken fluency and cohesion and coherence.

Table 5-6 shows the CEFR descriptors for the common reference levels provided under the heading 'Coherence and Cohesion'.

**Table 5-6. The CEFR scale of coherence and cohesion. Source: Council of Europe 2001: 125, emphasis mine, where 1) bold, red text denotes words related to prescriptivism, and 2) bold green text highlights words that signal either writing or speaking, which alternate in the scale, suggesting an inconsistent focus.**

Coherence and Cohesion	
C2	Can create coherent and cohesive <b>text</b> making full and appropriate use of <b>a variety of organizational patterns</b> and <b>a wide range of cohesive devices</b> .
C1	Can produce clear, <b>smoothly flowing</b> , well-structured <b>speech</b> , showing <b>controlled</b> use of organizational patterns, connectors and cohesive devices.
B2	Can use a variety of linking words efficiently to mark clearly the relationships between ideas.
	Can use a limited number of cohesive devices to link his/her <b>utterances</b> into clear, coherent discourse, though there may be some <b>'jumpiness'</b> in a long contribution.
B1	Can link a series of shorter, discrete simple elements into a connected, linear sequence of points.
A2	Can use the most frequently occurring connectors to link simple <b>sentences</b> in order to tell a story or describe something as a simple list of points.
	Can link groups of words with simple connectors like 'and', 'but' and 'because'.
A1	Can link words or groups of words with very basic linear connectors like 'and' or 'then'.

The other scale of interest, spoken fluency, is presented in the next table.

**Table 5-7. The CEFR scale of spoken fluency. Source: Council of Europe 2001: 129, emphasis mine, where bold, red text denotes words related to prescriptivism and NS language ideologies.**

Spoken Fluency	
C2	Can express him/herself at length with a <b>natural effortless, unhesitating</b> flow. <b>Pauses</b> 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 <b>natural, smooth flow</b> 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 <b>hesitant</b> as he/she searches for patterns and expressions, there are few noticeably long <b>pauses</b> . Can interact with a degree of fluency and spontaneity that makes <b>regular interaction with native speakers</b> quite possible <b>without imposing strain</b> on either party.
B1	Can express him/herself with relative ease. Despite some problems with formulation resulting in <b>pauses</b> and <b>'cul-de-sacs'</b> , he/she is able to keep going effectively without help.
	Can keep going comprehensibly, even though <b>pausing</b> 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 <b>pauses, false starts</b> and reformulation are very evident.
	Can construct phrases on familiar topics with sufficient ease to handle short exchanges, despite very noticeable <b>hesitation</b> and <b>false starts</b> .
A1	Can manage very short, isolated, mainly <b>pre-packaged</b> utterances, with much <b>pausing</b> to search for expressions, to articulate less familiar words, and to repair communication.

The descriptors in Tables 5.6 and 5.7 give rise to a number of essential questions: Since grammars of spoken language do not exist, what ideology is behind these

descriptors? Are the descriptors indicative of ideologies based on formal written language? Are the characteristics common to planned or unplanned speaking? To formal and/or informal spoken language? Is the focus of the scale clearly on either writing or speaking? With these questions as the guide, I will discuss the language ideologies guiding the assessment of pragmatic competence in these two CEFR scales.

In the coherence-cohesion scale in Table 5-6, an examination of the key words in each descriptor indicates that the scale makes no distinction between formal written language and spoken language. For instance, the level C2 descriptor mentions ‘text’ and A2 ‘sentences’, while B2 uses ‘utterances’ and C1 ‘speech’. Another key word in the C1 descriptor is ‘control’. This lack of distinction signals prescriptivism as the basis of the assessment, which points to standard language ideologies as the guiding principle. Previous research into the pragmatic use of cohesive devices also indicates formal spoken language to include a narrow range of connectors as opposed to “a wide range of cohesive devices” (Europe 2001: 125). For instance, Biber (2006) found two major types of discourse markers in a corpus-based study on university language: 1) discourse markers for spoken register, including *ok*, *so*, *well*, and *now*, and 2) linking adverbials for both spoken and written registers that included *however*, *therefore*, and *for example*. Among the discourse markers found to be common to all spoken university registers, *now* was the most commonly used to initiate a new topic and *ok* could be used interchangeably for this function. He also found *so* and *ok* to be the most common discourse markers, and used ten times more frequently in spoken academic registers than in written ones. In contrast, the linking adverbials *therefore*, *for example*, and *that is* were used only occasionally in lecturing. Thus, it appears that a limited number of cohesive devices for linking utterances during an academic lecture may very well be sufficient, for example, level A2. What this difference illustrates is that the formal rules of prescriptive language, rooted in standard language ideologies, are not relevant to the assessment of pragmatic competence in spoken language.

We know that spoken language differs from written language in two primary ways: it is unplanned and it happens in real time. Given these dimensions of spoken language, Biber et al (1999: 1066) point out that self-correction and reformulation can only occur through hesitations, false starts, and other dysfluencies since spoken language is constructed and interpreted under real-time pressure. Here, it is important to note that their research discusses spoken language, not learner language. These differences give rise to the question: Could it be that dysfluencies serve a purpose for both natives and non-natives in the real-time processing of language? There is some research that supports this view. According to Field, “a speaker who hesitates and inserts a lot of pauses assists the non-native listener because they mark proportionately more word beginnings and endings” (2003: 36). On pausing, he suggests four important functions:

- 1) At the end of a clause, to remove what is in our speech buffer (i.e. the group of words just produced) and to replace it with a new chunk of speech
- 2) Speaker finds difficulty in retrieving an item from the lexicon
- 3) At the end of an utterance, pauses may indicate that the speaker is prepared to hand over the turn to the listener

- 4) Pausing can also be used rhetorically to indicate that what comes next is of importance

(Field 2003)

Thus, it appears that hesitations, false starts, and dysfluencies are tools that help listeners and speakers to keep track of what is being said whether they are NSs or NNSs. Yet, hesitations and pausing are mentioned throughout the CEFR scale on fluency in Table 5-7 in a negative light: The descriptions seem to imply that hesitation and dysfluencies are both unusual and unnatural for NSs – although such features are an everyday part of conversation in NS-NS interaction. This unnaturalness also implies the presence of prescriptivism, and thus assessment criteria rooted in standard language ideologies.

In addition, examination of NS unplanned spoken language reveals that it consists of shorter chunks (Mauranen 2007, lecture series), instead of discourse spoken “at length with a natural effortless, unhesitating flow” (Europe 2001: 129). Unplanned, spoken conversation consists of incomplete structures, such as ‘*it was a bit erm...*’ as well as many very small units, such as *right, yeah, ok, looking good, and just a bit* (Mauranen 2007, lecture series). In addition, the clause appears to be a more relevant unit than the sentence since utterances as sentences can be hard to identify. Moreover, subordinate clauses are not always clearly connected to any particular main clause or then the relation may seem unusual (ibid). All of these examples point to shorter chunks of discourse within spoken language as opposed to ‘lengthy’ bits. These differences also imply the prescriptive rules of written language as the guiding principle for assessing spoken language.

Although there may be times when a speaker pauses to reflect on word choice, research has shown spoken language to be both inexplicit and vague. It is generally agreed that conversation is particularly inexplicit. Sinclair states that “it is a characteristic of a competent discourse to be inexplicit” (Sinclair quoted in Warren 2006). Warren points out that inexplicit language is highly context-dependent, just as face-to-face conversation is. Moreover, different levels of explicitness depend on discourse type. One form of inexplicitness is deictic references, such as ‘*what’s this box here?*’ Deictic expressions are unspecific and sometimes hard to understand when context-independent. They, however, become specific in the particular context in which they are used, and signal to the hearer that interpretation of their meaning relies on the context (Warren 2006). For instance, a speaker may point to something or draw an illustration and say ‘*there, that’s what I mean*’ rather than pausing to reflect on the precise word. Substitution and reference are still other forms of inexplicitness in spoken language. In discourse, inexplicitness features many grammatical words in proportion to lexical words (Mauranen 2007, lecture series). It is also most prominent when speakers have access to the same information (ibid). What previous research on inexplicitness suggests is that a speaker is more likely to fill lexical gaps with such expressions than to “[p]aus[e] to reflect on precisely the right words to express his/her thoughts” (Europe 2001: 129).

Like inexplicitness, vagueness in spoken language is common. Some typical expressions include the following: *thing, like, something, anything, whatever, kind of, sort*

*of, more or less, and stuff (like that)* (Mauranen 2007, lecture series). Previous research on vague language emphasizes its importance in conveying interpersonal meaning (Carter and McCarthy 2006; Channell 1994; Overstreet and Yule 1999). Its usage strongly indicates an assumed shared knowledge and can mark in-group membership. However, vagueness is often wrongly conceived as a sign of careless thinking or sloppy expression (Mauranen 2007, lecture series). Conversely, vagueness can function as a softener so that expressions do not appear too direct or overtly authoritative or assertive (ibid). Thus, speakers use them, particularly in informal contexts, where they prefer to convey information in a softer way.

Channell (1994) also argues that vagueness can create an informal and friendly atmosphere, express politeness and sometimes add a humorous effect. Moreover, it may be used in informative discourse when the speaker wants to give the right amount of information, appropriate to the communicative situation at hand, but not burden the listener with redundant and unnecessary information that would only obstruct the message (ibid). In addition, vagueness serves epistemic functions in academic speech. It characterizes a referent to serve the purpose for the moment but not to be taken as conclusive, for example *'here's a kind of a telescope by Buckminster Fuller'*. In this example, 'kind of' precedes an ad hoc descriptive label where the item under discussion is somehow not a typical telescope. In short, Channell (1994) divides the functions of vague language into two groups: those concerning information and those concerning interaction.

What previous research on spoken language indicates is that inexplicitness and vagueness, not preciseness (see level C2 of Table 5-7), are common features and they serve a communicative function. It thus appears that they would be better descriptors of spoken language than preciseness. Since preciseness does not reflect spoken language use, it appears that this CEFR scale is prescriptive as it hints at accuracy, a criterion for assessing written language.

## 5.4 Conclusion

Although the CEFR is a general framework for foreign language education, which aspires to be comprehensive yet open, dynamic and non-dogmatic (Europe 2001: 8), some of the descriptors on spoken language appear to be geared more towards written language that carry tones of prescriptive ideologies. In this same spirit, the CEFR discourse centers on an NS-NNS axis. Thus, a shift in focus from prescriptive ideologies of 'correctness' to a descriptive basis where language is described as it is actually used would be required in order to make room for variation in language use. It seems to me that the CEFR is more likely to achieve its goal of openness by relinquishing the evasive goal of NS competence and by welcoming the goal of a competent international educated speaker whose competence draws on plurilingualism.

## 6 Research design and methods

This chapter describes the methodology chosen to address the research questions presented in the introduction. It begins with the rationale for the research design, which guided the research approach. Following this, the chapter introduces the participants and instruments, after which the data collection processes are outlined and the analytical approaches described. In the last section, the reliability and validity of the research is discussed.

### 6.1 Rationale for the mixed methods approach

One of the backbones to good research lies in its design. In a mixed methods study, a good design will combine qualitative and quantitative methods in a way that “each highlights ‘reality’ [emphasis original] in a different, yet complementary way” (Lazaraton 2005: 219). Moreover, the practice of combining mixed methods in applied linguistics studies suits the field quite well since applied linguists are interested in studying simultaneously “both the exact nature (i.e. QUAL) and the distribution (i.e. QUAN) of a phenomenon”(Dörnyei 2007: 45). By applying both methods, it is possible to obtain information about both the micro- and macro-level of a given social context.

In adopting a mixed methods approach, this study employed an explorative strategy that involved a concurrent design. The two methods, used in a parallel manner, integrated the results in the interpretation phase of the study. The main purpose of this design was to provide a general picture through micro- and macro-level analyses, using data from two levels: the individual and the classroom. The approach combined self-reporting (from the individual lecturers) with observational data (from the students). The primarily qualitative data (from lecturers) was supplemented by questionnaire data (from students).

Giving consideration to the context of EMI in a university that houses a wide variety of multilingua-cultural backgrounds, my intention was to investigate professional English in this international setting. By using indirect methods, such as interviews, questionnaires, and stimulated recall, I hoped to capture some (not all) qualities of competence that could be used towards the development of criteria for assessing the professional English of EMI lecturers.

To investigate the student perspective, quantitative research was the most appropriate method. This scientific view provided evidence on whether there was an effect for time or theme in students’ perceptions of EMI-lecturer language competence. In this study, quantitative research provided a basis for capturing the realities of students’ perceptions of their lecturer’s English by means of an initial and post survey (Cohen et al. 2007). Pencil-and-paper questionnaires were favored since they are easy to administer and respondents can complete them without assistance, which minimizes the researcher effect (e.g. Bryman 2004; Dörnyei 2003a; Sale et al. 2002). The accuracy of this data was important since “any interpretation of data is only as good as the accuracy of those data” (Newman and

Benz 1998: 110). In this study, the rationale for including the student perspective is two-fold: 1) to collect detailed information on both perceived lecture comprehension and lecturer language use, and 2) to include a macro-level view that complements and confirms lecturer's self-perceptions of professional English.

On two-way communication, lecturing in this study involves a face-to-face domain, which is viewed as a social event with its own cultural rules that evolve around spoken language. In a lecture event, the presence of a lecturer and students is a social context where communication occurs through interaction. A series of papers by Schegloff (1982; 1988; 1995), all of which have the title "Discourse as an Interactional Achievement", spells out this position clearly. For instance, he stated:

It is some 15 years now since Charles Goodwin . . . gave a convincing demonstration of how the final form of a sentence in ordinary conversation had to be understood as an interactional product . . . Goodwin's account . . . serves . . . as a compelling call for the inclusion of the hearer in what were purported to be the speaker's processes.

(Schegloff 1995: 192)

This research adopts an interactional view to spoken language. This view perceives lecturing as an interactive performance that involves speakers and hearers. Following Hymes' notion of communicative competence, lecturing as an interactional activity necessitates communication that is suitable to the lecture environment. This interactive view contrasts with lecturing viewed as a simple, unidirectional projection of information. An interactional view also suggests that communicative competence cannot reside in a single individual.

This approach aligns with qualitative studies in postmodern research, where human behavior is dependent on both the context and cultural patterns. This dependency regulates communication among the social actors. At the micro-level, qualitative research provides methods for the interpretation of the social world through the eyes of the participants as reflected in their beliefs. At the macro-level, the classroom perspective is represented through the eyes of the students as participants in the lectures, and as valued clients of the university. Taking a mixed methods approach provides the link to the interactive aspect of lecturing, connected through micro- and macro-level analyses.

Taking a pragmatic position, my research was instigated by a problem that I identified: the lack of criteria for assessing the professional English of EMI lecturers. The intention guiding my research was to gather information from the lecturers' voices related to: 1) the ideological basis of EMI-lecturer language competence (Chapters 3 and 7), and 2) the potential of five selected CEFR scales to measure the professional English of EMI lecturers (Chapters 5 and 8). I argue that lecturers' accounts reflect their perceptions of ELF language competence for professional purposes. Their accounts include the socially constructed meanings accepted by them. This argument is relevant since I intended to explore the individual views guiding EMI-lecturer perceptions of their English for professional use.



## 6.2 Participants

This section presents an overview of the participants in this study. It first introduces the EMI lecturers, and then their students.

### 6.2.1 Overview of primary and secondary participants

All seven EMI lecturers have engineering backgrounds, and have volunteered to teach their subjects in English. In addition, they were participating in a pilot-mentoring program on teaching through English in a multicultural environment at Aalto University. Of these seven (A-G), four were primary participants (A-D) and three secondary (E-G). While the primary participants took part in the full study, the secondary participants participated in one interview on five selected CEFR scales. Table 6-1 presents an overview of the seven lecturers.

**Table 6-1. Overview of primary (A-D) and secondary (E-G) participants.**

	A	B	C	D	E	F	G
<b>Language background</b>	L1: Bilingual Arabic/French L2 English L3 Finnish	L1 Finnish L2 English L3 Swedish	L1 Finnish L2 German L3 Swedish L4 English	L1 Dutch L2 French L3 English L4 Finnish	L1 German L2 English L3 French L4 Finnish	L1 Spanish L2 English L3 Finnish	L1 French L2 English L3 Spanish L4 Finnish
<b>Years of English at school</b>	3	10	3	6	9	4	12
<b>Years working in English</b>	13	12	10	32	16	20	18
<b>Years lecturing in English</b>	13	12	1 ½	8	10	4	7
<b>Age range</b>	41-50	41-50	51-60	51-60	41-50	41-50	41-50
<b>Number of students</b>	8	6	17	5	N/A	N/A	N/A

As shown in Table 6-1, the seven lecturers have quite different backgrounds, even though two have Finnish as their L1. They are all multilingual and the number of years they studied English at school varies widely. In spite of this, they all have lengthy experience using English at work, either in an academic or business environment, or then both. The age range suggests a mature group. As noted in the last row, the class size tends to be small as this is preferred in the university for master-level courses. Only C's class exceeds the recommended limit of ten to twelve students. Because all seven lecturers volunteered to lecture in English and to participate in a pilot-mentoring program, it suggests that they are all fairly confident with their skills in English and open to new experiences.

## 6.2.2 Primary participants

This section provides detailed descriptions of the four primary lecturers participating in the study. Table 6-2 presents an overview of A-D’s self-assessments of their English on a 4-point Likert scale, obtained during Interview 1. Lecturers (A-D) each rated their English proficiency for listening, speaking, writing, and lexicon. The self-assessment is based on lecturers’ perceptions of their English proficiency in two contexts: 1) everyday use, and 2) professional use. The purpose of this self-assessment was to ascertain lecturers’ perceptions of their English skills in order to determine what qualities, if any, they felt needed improvement.

**Table 6-2. Self-assessed English for professional and non-professional purposes.**

Lecturer A					Lecturer B					Lecturer C					Lecturer D				
	Less than adequate	Adequate	Good	Excellent		Less than adequate	Adequate	Good	Excellent		Less than adequate	Adequate	Good	Excellent		Less than adequate	Adequate	Good	Excellent
<i>English for non-professional purposes</i>					<i>English for non-professional purposes</i>					<i>English for non-professional purposes</i>					<i>English for non-professional purposes</i>				
Listening		x			Listening				x	Listening			x		Listening				x
Speaking	x				Speaking			x		Speaking		x			Speaking				x
Writing		x			Writing			x		Writing			x		Writing				x
Vocabulary	x				Vocabulary			x		Vocabulary			x		Vocabulary				x
<i>English for professional purposes</i>					<i>English for professional purposes</i>					<i>English for professional purposes</i>					<i>English for professional purposes</i>				
Listening			x		Listening				x	Listening			x		Listening				x
Speaking		x			Speaking			x		Speaking			x		Speaking				x
Writing		x			Writing			x		Writing			x		Writing				x
Subject-specific terminology			x		Subject-specific terminology				x	Subject-specific terminology				x	Subject-specific terminology				x

Looking across all four tables, it is clear that B and D rated their proficiency levels higher than A and C did. Only A and C have marked one or more skill areas as ‘adequate’ or ‘less than adequate’ while B and D have marked all skill areas as ‘good’ or ‘excellent’.

### 6.2.2.1 Lecturer A: L1 Arabic and French

Although A had only three years of English (L2) at school, he started his tertiary studies in English in 1991 upon arriving in Finland. As a self-directed learner, he studied English for one year and then passed the TOEFL test. After his studies, A remained in Finland and has continued to use his English for professional purposes. Because he actively uses his professional English, he believes that it has developed over the years and that it is adequate for lecturing on engineering topics. He reported his knowledge of field-specific terminology to be strong, but described being able to pronounce words correctly as another matter. He views fluency as a potential problem, but one that he can accommodate

through the use of his acquired linguistic range of English and French. Thus, he utilizes grammatical structures from French if he cannot find a structure he would need in English. He also overcomes missing words or expressions by finding alternative ways to express himself. As a way to further accommodate his NNS students, he avoids using idiomatic expressions. For him, the most important point is to communicate his ideas clearly to his students. He believes that he can best enhance communication by speaking with correct grammar. Because he views his pronunciation as his main weakness, he tries to compensate listeners by speaking as grammatically correct as possible. This means following the prescriptive rules of ELT grammar books. However, A seems to have conflicting views about the NS. While he would like for his pronunciation to be more native-like, he disparages ENL speakers whom he views as unaccommodating in NS-NNS interaction. On native-like pronunciation targets, he seems to believe having pronunciation that approximates native-likeness will ease ELF communication by reducing variety in accents. He also mentioned that NSs notice mistakes made by NNSs and even judge NNSs according to whether they speak correctly. For this reason, he feels it is important to speak as correctly as possible even though there may be no ENL speakers in his lectures.

#### *6.2.2.2 Lecturer B: L1 Finnish*

Having had ten years of English (L2) at school, B has strong skills in English although she has never lived abroad. She reported that she could lecture equally well in both English and Finnish. Although she reported that she lectures fluently in English, she requested help in finding ways to both maintain and increase her fluency in English. On professional English, B reported having excellent knowledge of field-specific terminology. She attributes this to working with systems that are in English only, for which she also received training by ENL speakers. On other aspects of her English, she reported having strong grammar skills and being able to speak quite correctly. She believes it is important to speak as grammatically correct as possible, which carries a prescriptive tone. She also believes her intonation to be dull to the extent that students have difficulty staying awake in early classes. Her desire would be to attain a British accent, although she acknowledges that she would not need a near native accent to be intelligible in her lectures. During lectures, B mentioned that she finds it quite challenging to understand some students who have strong Asian accents in English. Viewing her English as better than NNS students, B feels confident about her English and believes that she is a good model for them (as a professional working in English). Conversely, she feels less confident when she knows there are ENL speakers present in her lectures.

#### *6.2.2.3 Lecturer C: L1 Finnish*

Like A, C studied English at school for three years. Although he found language learning at school to be difficult, he became highly motivated as an adult to improve his English (L4) while working in an international, multicultural corporate setting for ten years, two of

which were spent in the USA. After leaving the corporate world, he continued to use his English at work, and now in an academic environment as a lecturer. Although he has been actively using his English, he feels that his spoken fluency is lacking. He also mentioned that it can be difficult to understand different accents, especially Asian ones, and different speakers, who for example speak too fast irrespective of being ENL speakers. He feels that it would be nice to speak English correctly, but added that it is not the most important point. Nor does he target speaking like a NS of English. He does not feel that this would be necessary to be intelligible in EMI lectures. For him, the most important point is to express his ideas clearly. C finds it much easier to communicate in English in Finland among SLUs than to participate in meetings dominated by NSs, for example in the USA. In Finland, C feels comfortable having English with imperfect grammar and a Finnish accent, but uncomfortable in the USA or England where he feels pressure to conform to local standards. In EMI lectures, C's focus is on engaging his students rather than on speaking 'perfect' English, which seems to enhance his self-perception and confidence.

#### *6.2.2.4 Lecturer D: L1 Dutch*

D studied English (L3) at school for six years, moved to Finland in 1979, and began using his English for professional purposes – initially for publishing and then later (2004) for teaching. He feels that he does not need help with his English; rather he believes that he is too fluent in English for his NNS students. He views his English as equivalent to a NS, and added that the main problem in EMI lectures is that he knows English too well. Consequently, he may speak too fast and even ramble when he gets excited about a topic, making it difficult for students to follow. For this reason, he tries to remember to speak a bit slower and to really try to formulate his sentences clearly. Because his proficiency level is lower in Finnish than in English, he believes that he delivers a better lecture in Finnish. On English language needs, D requested help with getting the message across better in a multicultural teaching environment. On this point, he added that strong accents can be a challenge, regardless of whether the speaker is native or non-native. However, he encounters strong accents less frequently in professional situations with other NNSs than when using English as a tourist. He attributes this difference to these professionals having acquired accommodation skills for communicating in international contexts and to being educated. For lecturing, he also said that he sometimes lacks vocabulary, for example, when he needs to explain a Finnish concept for which there may be no equivalent word in English. In this situation, he uses the Finnish word(s) and then explains the concept in English or illustrates it visually. On vocabulary, he also mentioned that he cannot use his full range since he believes that students have a smaller range than he has. Overall, D feels competent lecturing in English and described his English as being 'good enough'. He reported his lecture performance in English as being similar to a competent NS who delivers a slightly incompetent performance due to sentences not always being perfect, but who is perfect otherwise. He described imperfect sentences as being incomplete thoughts, repeating a word, hesitating when speaking, and other similar dysfluencies that appear in spoken language. In addition, he described his accent as having a Dutch fingerprint.

### 6.2.3 Students

The total number of students participating in the lectures was forty-two, of which thirty-six completed both the initial and post questionnaire. All students were participating in EMI lectures at the master’s level. Table 6-3 presents the first languages of the four lecturers and their students as well as students’ self-assessed English.

All students self-assessed their English on a 4-point Likert scale, ranging from ‘excellent’ to ‘poor’. On this self-assessed scale administered during the first class (in Appendix G), Table 6-3 shows that the majority (95%) of students assessed their language skills as ‘good’ or ‘excellent’. Only two students rated themselves as having ‘fair’ English, and none as ‘poor’. In general, their self-ratings suggest that they should have no real problems in following a lecture. There was no significant change in self-assessment in the post survey. In total, sixteen languages were represented in the study. On the whole, the groups were quite international, which means that English was a true lingua franca in each sample.

**Table 6-3. Overview of lecturers (A-D) and their students.**

Master’s level engineering: Lecturers A-D and their students			
	Lecturer L1s	Student L1s	Student self-assessed English
A	Arabic/French	English (1)	Good
		German (1)	Good
		Italian (1)	Good
		Korean (1)	Good
		Lithuanian (3)	Good
		Urdu (1)	Excellent
B	Finnish	Estonian (2)	Good
		French (3)	Good
		German (1)	Good
C	Finnish	Amharic (1)	Good
		Chinese (1)	Excellent
		Chinese (3)	Good
		Czech (1)	Fair
		English (1)	Excellent
		Finnish (3)	Excellent
		Finnish (4)	Good
		Nepalese (1)	Excellent
		Swedish (1)	Good
		Urdu (1)	Excellent
D	Dutch	Finnish (2)	Excellent
		Finnish (1)	Good
		Finnish (1)	Fair
		Finnish (1)	Good

### 6.3 Sources of data

The methods used in my fieldwork to enter and become familiar with the world of EMI lecturers were multiple. The sources of data collection were primarily qualitative, including interviews, stimulated recall, self-assessment, and a field journal. These data were supplemented with a quantitative survey involving repeated measures. The choice of methods was based on three considerations. First, the naturalness of the instruments in relation to my role as mentor: they were the tools that helped me to help my mentees (i.e. the EMI lecturers). Second, lecturing is an interactive event, which necessitates examining the perceptions of both parties involved, that is, the lecturer and their students, in order to establish face validity. A third consideration was related to tackling the qualitative issues raised in the research questions while adding strength to the trustworthiness of the study through triangulated methods ( see Dörnyei 2007).

### 6.3.1 Triangulation strategies

The instruments used for triangulation each had their own focus and shed light on different aspects of professional English used in an international context. Thus, triangulation in this study was not a matter of establishing whether the data analyses from each instrument would lead to the same results (Gliner 1994). Rather, the data from the different instruments were combined to develop a comprehensive view of the perceptions of EMI-lecturer language competence. With two levels of data (i.e. individual and classroom), this approach allows for the combination of micro and macro perspectives, with quantitative research tapping the macro-level perceptions of EMI lecture comprehension and qualitative analysis tapping the micro-level perceptions of professional English as viewed by individuals. Table 6-4 presents the triangulation strategy at a glance.

**Table 6-4. Data and method triangulation strategies.**

Stages	Research focus	Instruments	Steps in analyses	Outcomes
1	RQ1, H1 Language ideologies informing positive and negative self-assessments	Interview 1 & 2 Stimulated recall	1. Per lecturer (A-D), search for broad similarities in the data and establish preliminary themes 2. Across lecturers (A-D), combine sets of related beliefs and construct a generalized version of the accounts under one theme	From individual perspective: Patterns in EMI lecturer language ideologies that relate to positive and negative self-perceptions of professional English
2	RQ2, H2 Suitability of five pre-selected CEFR descriptors for assessing EMI lecturers of engineering	Interview 3	1. Establish broad similarities and preliminary themes from lecturer's (A-G) accounts 2. Combine related sets of beliefs and construct a generalized version of the accounts under one theme	From individual perspective: Patterns in the CEFR ideological and theoretical basis in relation to EMI lecturers of engineering
3	RQ3, H3 Changes in students' perceptions over time of their lecturer's spoken English	Repeated survey	1. Analyze three major themes in repeated survey with ANOVA: Language use, communication, and lecture comprehension 2. Analyze qualitative data from the repeated survey	From students' perspective: Whether students' adjusted to EMI-lecturer spoken language over time
4	Combine the results from the above stages to present a general picture of the findings in relation to the development of assessment criteria for spoken professional English in an international context			

In Table 6-4, the first and second stage of my triangulation strategy involved qualitative data that was digitally recorded. Prior to recording, consent forms were signed by each lecturer (A-G). These two stages represent the micro-level analyses in the present study. The third stage included primarily quantitative data collected anonymously from the students for macro-level analysis. This analysis, however, also included qualitative data from the open-ended question in the survey. This mixed methods design relies on qualitative and quantitative data for making meaning out of participants' responses, as explained in the rationale. In the planning stage, differentiating the research questions aided in identifying the types of data to be collected and the appropriate instruments. The results from the three stages were then combined to present a general view of professional English in an international context. The steps in the data analyses are described in section 6.4.

This section will next present the different sources of data in the present study. Although several sources of data were utilized, due to space limitations, this section will concentrate on describing the three main sources of data, including interviews, stimulated

recall, and a repeated survey. It then briefly mentions other sources of data that are primarily backgrounded in the study.

### 6.3.2 Interviews

The main research instrument in the present study was semi-structured interviews. This method is regularly employed in qualitative research, and is a suitable approach for trying to determine how lecturers (A-G) perceived their professional English for working in an international context.

Of the seven lecturers, the primary lecturers (A-D) participated in three semi-structured interviews and the secondary lecturers (E-G) in one semi-structured interview. The two interviews involving only the primary lecturers (A-D) included the following: *Interview 1* was conducted at the onset of the mentoring sessions. The purpose of the first interview was to determine each lecturer's language background and his/her EMI-lecturer needs (1 hr). Prior to the interview, a self-assessment form (in Appendix A) had been completed via email, and this self-assessment served as the prompt for questions in this interview. The interview protocol can be found in Appendix B. *Interview 2* was conducted immediately following a lecture observation to discuss lecturers' perceptions of their teaching performance (2 x 1.5 hrs)<sup>15</sup>. For Interview 2, the data included hand-written field notes taken during the lecture observation, a video-recorded lecture, and an audio-recording of the interview (1.5 hrs). The interview protocol is presented in Appendix C. All seven lecturers (A-G) participated in *Interview 3*, which involved individual interviews, each of which was digitally recorded (1.5 hrs). This interview was conducted after the pilot-mentoring certification assessment had been completed. It covered the five CEFR scales that had been used in that assessment. All lecturers were familiar with the scales, which had been emailed to them with the task instructions for the pilot-certification assessment. The purpose of this interview was to ascertain lecturers' views of five selected scales as potential assessment criteria for the purpose of lecturing in EMI. In the interview, the CEFR scales served as a prompt for discussion. The interview protocol can be found in Appendix D.

As this study involves self-perceptions of EMI-lecturer experiences in lecturing as well as their judgments and self-reporting methods, interviews were favored since they "allow researchers to investigate phenomena that are not directly observable, such as learners' self-reported perceptions" (Mackey and Gass 2005: 173). Although semi-structured interviews are based on prepared interview guidelines, the interview questions are characteristically open-ended, which allows room for interviewees to elaborate, contrary to structured interviews (Dörnyei 2007: 135-136). The semi-structured interviews produced rich data for the analysis, as expected. The complexity of the data also reflects the complex realities of the individuals (ibid: 125).

---

<sup>15</sup> Although I observed two lectures, the first one was followed by an interview and the second, which was videoed, served as data for the stimulated recall interview.

In conducting an interview, the role of the interviewer is challenging. It involves a balancing act between being neutral and providing sufficient input to the interviewee. The term 'inter-view' consists of two parts, where 'inter' means *between* and 'view' means *a way of considering something*, which together suggest co-constructing views in a social exchange. Thus, part of the balancing act in interviews includes reducing potential bias. For the interviewer, this means maintaining a neutral position while creating space for the interviewees to elaborate freely (ibid: 141), as well as taking measures to reduce a 'social desirability bias'. Such measures include establishing good rapport, not presenting an image of being perfect, or wording questions to suggest that some behaviors are common (e.g. 'Many teachers now think that ...') (ibid). However, the literature also presents an alternative position on neutrality. For instance, Fontana and Frey (2005) argue that co-constructed speech involves a social exchange where taking a stance is unavoidable. They advocate 'empathetic interviewing', and claim that it elicits a more honest response since the interviewer is viewed as an ally. Both alternatives make relevant points, where it seems that a proper balance between neutrality and empathy could be a good strategy. In the interview, offering options to the interviewee and encouraging elaboration can also result in richer data (ibid: 142-143). Simultaneously, the role of the interviewer includes minimizing his/her contribution (Cohen et al. 2007: 362). As an interviewer, I tried to keep these objectives in mind and to employ them in the interviews.

The primary tool in a semi-structured interview is the interview guide. A strength in this tool is the flexible, yet systematic approach that it allows (Dörnyei 2007: 143). Using the guide, an interviewer can vary the order of the questions, thereby creating a more natural flow in the conversation. This explanation also describes how this tool was applied in the present study.

Following the recommendations in the literature for conducting effective interviews, I tried to make the lecturers as comfortable as possible. This included conducting the interviews in a familiar milieu, such as the lecturers' offices, when possible, and otherwise in a room they had selected. To create a friendly, relaxed environment, I started the interviews with small talk in an effort to encourage the lecturers to speak as freely as possible. During interviews, I also mirrored lecturers' responses by repeating some information that would help them to reflect and provide input (Mackey and Gass 2005: 174-175).

My role as the interviewer was to facilitate and guide, not to dictate exactly what would happen during the interview. Thus, I used the interview guide to indicate the general area of interest and to furnish cues when needed. This procedure allowed the respondents a role in determining how the interview would proceed. I tried to be a good listener and to use minimal probes (ibid: 66). To explore lecturers' views, both flexibly and in detail, I first introduced broad open questions, asking lecturers to describe their experiences. The prepared questions guided the discussion, but were not strictly followed since it was more important to follow the lecturers' lines of thought. If the conversation veered to a different direction, then I followed their lead in order to hear what they found to be interesting. This means I randomly chose questions from my interview guide, when needed, in order to create a more natural flow in the interview. Because I was interested in obtaining responses to both general and specific views, my interview technique also



employed, at times, a funneling technique (Smith and Osborn 2008: 62), illustrated in Table 6-5.

This technique allows the respondents to first give their general view before being funneled into more specific questions of particular interest to the researcher (ibid). Furthermore, in the event that the lecturer's first answer also addressed the targeted issue, then I did not redundantly ask more specific questions. An interview conducted in a reverse sequence would more likely produce a bias in favor of the researcher's specific concerns (ibid). Even though the interviews were recorded, I also took notes that helped me to immediately plan the ongoing choice of questions.

Although semi-structured interviews produced useful data for the present study, they have also been criticized for producing poor data. One reason for this view is the lack of standardization in the interviews. Asking different questions to each interviewee, as opposed to the exact same ones, can produce very different data, making it

**Table 6-5 Examples of funneling technique.**

Funneling technique	
1.	How would you describe your accent in English?
2.	How do you feel about your accent? Or, do you like your accent in English?
3.	(If negative, another possible question) Are you concerned about other people recognizing your accent as being influenced by your L1?
.....	
1.	Where would you place yourself on this scale for professional purposes? for non-professional purposes? Any differences?
2.	What guided you in your choice? Or, what key words caught your attention?
3.	Can you pinpoint any specific words or phrases?
4.	What do these words or phrases say to you?

difficult to compare the responses. This caveat, however, is outweighed in the strength of this type of research. The non-standardized approach is not intended to be repeatable. Rather, it allows flexibility in exploring a complex, dynamic situation, for which the data reflects the reality at that point in time. Thus, an attempt to replicate this non-standardized research would not be realistic without undermining the strength of this type of research.

### 6.3.3 Stimulated recall

Another instrument for eliciting data in the present study involved stimulated recall (Gass and Mackey 2000), which centered on forty-five minute lectures given by lecturers (A, B, and D), each of whom were video recorded. The use of stimulated recall would allow lecturers to reflect on their thinking during lectures. Because this method uses video to capture the participants in action, it can be used to recreate the original situation, thus allowing lecturers to better reflect on their lecture experiences in the specific situation (Calderhead 1981).

Stimulated recall is a method widely used to gain qualitative insight into working memory processes (Gass and Mackey 2000), and has been used in both cognitive and non-cognitive studies. For instance, it has been used as a tool to investigate individuals' perspectives on learning (Erickson and Mohatt 1977), to study children's reasoning abilities (Hample 1984), and to research questions related to teachers' actions in the classroom (Calderhead 1981a; 1981b). It has also been used in SLL research, for example, in oral interaction literature (Dörnyei and Kormos 1998; Tyler 1995) and in interlanguage

pragmatics literature (Cohen and Hosenfeld 1981; Faerch and Kasper 1987). Furthermore, it is a method that is often employed in connection with other methodologies.

Like every method, stimulated recall has its pluses and minuses. On the plus side, the multimedia resource provides cues that were present while performing the task (Mackey and Gass 2000, 2005). The research is also performed in a natural setting, representing a real-life context. It works as a valuable tool when the recall session occurs as soon after the event as possible, making it less likely for participants to rely on memory alone (ibid). Thus, the advantage of stimulated recall over a post interview is that the former uses an aid to stimulate the memory whereas the latter relies heavily on memory without prompts to assist. As a research tool, it also allows relatively unstructured responses from the participants. Mackey and Gass (2005) suggest that it is an effective way to gain insight into participants' interpretation of events and their thoughts at a particular point in time. On the minus side, participants may not have the language needed to express deeper thoughts in a foreign language (Gass and Mackey 2000). Another minus is that participants may censor their thoughts in an effort to present themselves in a more positive light (ibid). Nor does stimulated recall capture actions over time or the complexities of classroom interaction (Lyle 2003). For these reasons, other methods were also used to collect data related to assessing the professional English of EMI lecturers, including observations and interviews.

The aim of this methodological choice was to provide the lecturers with an opportunity to express their thoughts about their interactive processes while lecturing in ELF. The videoed lecture provided a stimulus for EMI lecturers to talk about their spoken performance concerning lecturing in English in relation to the immediate context of their classrooms. It is acknowledged, however, that a danger with this method is that it may have led lecturers to simply provide post hoc rationalizations. In the data collection, the researcher's role was made clear to the participants. Although an assistant from the pilot-mentoring program recorded the video, I was present during the session. Lecturers themselves had suggested the video date for collecting the data. The data gathered included researcher observation, video-recorded lectures, and participant stimulated recall interviews that were audio-recorded.

Due to my lack of experience in working with stimulated recall, I followed Gass and Mackey's (2000) recommendations and modeled my interview protocol on theirs, the latter of which can be found in Appendix E. In the end, however, the interview protocol (in E) was not closely followed.

According to Gass and Mackey, controlling the timing between the video and recall events is important to the accuracy of the recall (ibid). Thus, they recommend conducting the introspective interview as soon as possible after the video session, and preferably within two weeks (2000: 54-55). The videos for lecturers A, B and D were available for viewing within two days of their video sessions, and were accessed via Optima, an online management platform. For C, there was no video session as he had chosen a different alternative as part of his training in the pilot-mentoring program<sup>16</sup>. In the best case (A and

---

<sup>16</sup> Because participating in a video session was optional in the pilot-mentoring program, C chose the alternative of being observed without video and participating in a reflective follow-up interview.

D), the stimulated recall interview occurred on the second day, and in the worst case it was rescheduled to one week later (B).

In terms of approaches used with stimulated recall, different paths can be taken. While the focus of the present study was on the professional English of EMI-lecturers, it was not known which aspects of the videoed lecture might be important. For this reason, the entire 45-minute video was viewed in the recall session. As Gass and Mackey recommend training participants for the introspective interview using simple instructions, I followed this advice. I also took their advice on getting participants involved in selecting and controlling the stimulus episodes (in order to reduce the likelihood of researcher influence). Thus, for the training session, the participants were asked to view their own videos and to mark the time points for any aspects of interest on their teaching performance in English. This preparation allowed the lecturers to provide the structure for the stimulated recall session, which ultimately changed the interview protocol planned (in Appendix E). It seemed more sensible to follow their lead than to follow a strict protocol that I had not previously tried. Although I had prepared a few questions for the interview, my objective was not to guide or influence participants' analyses of their own performances. At the onset of the interview, the participants indicated that they had understood the instructions given for training, had viewed their videos, and were prepared for the recall session (1.5 hrs). Thus, we began and I let them take the lead and control the session, which was actually a good strategy since many of the points they selected (in their videos) would have passed unnoticed by me. Although I did not strictly follow my original plan for the interview protocol, the approach employed nevertheless provided data relevant to the study and with minimal interviewer influence.

#### **6.3.4 Repeated survey**

The third main instrument, the questionnaire, is one of the most common methods of data collection in SLL research (Dörnyei 2003b: 1). Ease of construction is cited as its main strength, but also as the main weakness: "questionnaires with sufficient (and well-documented) psychometric reliability and validity are not that easy to come by in our field" (ibid: 3). Thus, questionnaires need to be used with some caution. Their strengths include versatility and unparalleled efficiency as regards researcher time, effort, and cost (ibid). These features make it possible to acquire a vast amount of information in a short time on an assortment of topics from a variety of individuals and situations. Furthermore, if the questionnaire is well designed, processing the data can be straightforward and quickly achieved with statistical software. Questionnaires are also regularly used to measure factual, behavioral, and attitudinal data (ibid: 8), making it suitable for the present study. Apart from these virtues, questionnaires also have limitations, which can lead to unreliable and invalid data caused by ill-constructed items. Other limitations include low response rates, misunderstandings, self-deception, social desirability bias, and some other shortcomings (see Dörnyei 2003b). With these pluses and minuses in mind, the questionnaire was used for this study, with careful planning and piloting as measures for avoiding pitfalls.

A total of thirty-six students completed the initial and post questionnaire, out of forty-two participants. The repeated survey was administered to students in A-D's courses, at the end of the first and last lecture. The purpose of the repeated surveys was to collect information on change in students' perceptions of their lecture comprehension over the duration of a course (either 7 or 13 weeks in duration) to determine whether they adjusted to their lecturer's spoken English. Although previous studies have shown that EMI students of physics and engineering adapted to lecture comprehension over a one-year period (e.g. Airey 2009; Klaassen 2001), I have not found studies that examine EMI students' perceptions of English in lectures over a shorter period of time.

The questionnaire used in this study was an adapted and expanded version of one piloted in Suviniitty (2012). The adaptations were partly based on Suviniitty's suggested changes<sup>17</sup>, regarding items that had not worked well and items she felt had been missing. This discussion resulted in my keeping eight of the original items and adding twenty-four new ones. While the original questionnaire primarily covered items related to lecture comprehension, content, and classroom atmosphere, the additional items in the revised version covered terminology, grammar, pronunciation, utterance length, and coherence. These questions were generated from my pilot data, collected during observations and interviews with Lecturer A in fall 2009, as well as from discussions with my doctoral supervisor. The revised questionnaire was then presented in a research seminar to my supervisor and doctoral students, all of whom helped in piloting it. The primary changes were related to clarifying the wording and ordering of some items.

The questionnaire is presented in Appendix F. It consisted of four parts with a total of 36 items. Part 1 with 32 items was designed to elicit students' perceptions on three themes related to the lecturer's professional English: language use, communication skills, and lecture comprehension. Each of the randomly arranged 32 items was rated on a four-point Likert scale, comprised of 'agree', 'partially agree', 'partially disagree', and 'disagree'. Part 2 with two items focused on a comprehensive view of the lecturer's English. Following this set of questions, Part 3 contained one open-ended question for comments, after which personal language background information was collected.

To avoid obtaining responses on only one side of the questionnaire, the items were worded both positively and negatively. This approach helps to reduce possible bias (Dörnyei 2003b:106). Consequently, for scoring, the following items were reverse coded: Q2, Q4-Q6, Q8, Q15-Q18, Q20-Q23. In addition, the scale direction was corrected, where 1 = disagree and 4 = agree.

### 6.3.5 Other methods

Other instruments used in the analysis involve a *self-assessment questionnaire* (included in Appendix A), a *CEFR self-assessment* using five selected scales (in Appendix D) and my *field journal*. The hand-written notes in my journal have aided in writing the

---

<sup>17</sup> Note: This information was collected through personal discussion, as Jaana is a colleague in the same research group as well as working at the same university.

dissertation, but to a lesser extent than the recorded data. In addition to the interviews and stimulated recall, my fieldwork included *participant observation* in a number of events that helped to get acquainted with the professional culture of EMI lecturers as well as the training received in the pilot-mentoring program: mentoring seminars (2 x 8hrs), participating in mentoring workshops (3 x 3hrs), a mentors' circle (with twelve mentors; 10 hrs), observing classroom lectures (2 x 1.5 hrs), talking to students pre- and post-lecture, as well as interviewing the co-coordinator of the pilot-mentoring project (0.5 hrs). With regard to this dissertation, much of the information acquired through fieldwork involving participation in the mentoring seminars, workshops, and mentors' circle is largely backgrounded.

## **6.4 Analysis of the data**

This section describes the analytic procedures in the present study. It first presents the qualitative analysis, followed by the quantitative analysis.

### **6.4.1 Qualitative data analysis**

Taking a holistic approach to the data from interviews and stimulated recall, I transcribed the interview data in full as well as selected parts of the stimulated recall data. To answer RQ1, I combined the data from Interviews 1 and 2 with the stimulated recall data. In terms of combining sets of beliefs obtained from different lecturers (A-G), this procedure is in accordance with my definition in section 3.2, which describes language ideologies as shared as opposed to individual. These shared beliefs can also be both implicitly and explicitly stated. The combined interview (1 and 2) and stimulated recall data were thus the primary data for the analysis of language ideologies in Chapter 7. The broader, more widely known language ideologies to which lecturers' (A-D) language ideologies were then linked are discussed in the theoretical framework in Chapter 3. Similarly, to answer RQ2, I combined the data from lecturers (A-G) gathered in Interview 3. These data were the primary material for analyzing the adequacy of the CEFR descriptors for ELF and LSP.

To analyze the data, I listened to the recorded sources repeatedly, transcribing the data and taking notes simultaneously. This process helped me to get a sense of the whole. After transcribing, the coding began. The coding was a process of organizing the data into chunks before assigning meaning to the chunks (Creswell 2003: 192). To code the qualitative data, I followed the guidelines in Saldana (2009), who proposes a two-cycle method for analyzing data. As a researcher, my task was to construct a theory from this raw data.

In the first cycle, I applied descriptive coding to the transcribed data. The main aim of this cycle was to identify responses that seemed helpful in addressing the research question under investigation. During this cycle, it is necessary to read and listen to the transcripts repeatedly in order to become familiar with the data and to identify bits that

could contribute to answering the research question. In my analysis, I worked with the data, searching for broad similarities among the utterances. In my search, I looked for patterns related to language ideologies “*as conceptualized by my informants*” (Spradley 1979: 93, italics in original). When there were similarities that occurred frequently, I highlighted these accounts when they seemed helpful in understanding the participants’ experiences and perceptions related to the topic under investigation. For identifying meaningful chunks, the analytic procedures drew on discourse analysis as presented in Gilbert and Mulkay (2003[1984]). They also drew on advice from Jenkins (2007), who recommends paying attention not only to *what* informants say but also *how* they say it. In her work, Jenkins utilized a two-tier analysis that examined: 1) explicitly articulated beliefs, and 2) implied beliefs that represented latent attitudes as expressed through use of metaphor, repetition, and lexical choices (i.e. that displayed positive and negative language). Jenkins’ analyses also utilized prosodic features to contextualize cues, such as pauses, laughter, and stress. After identifying meaningful chunks, I grouped them thematically, and then relisted and regrouped them – revising on a continuous basis. The thematic codes were thus created through a data-driven approach. During the coding, I noted down brief, preliminary comments close to the highlighted data, which helped me to understand the specific chunks. My comments were then used to define the descriptive codes.

The second cycle involved interpretive coding. In this cycle, I bundled the descriptive codes and then defined them. This action meant going beyond the transcribed words, resulting in codes that focused more on my interpretation of the meanings of participants’ accounts. It also meant finding overarching themes that characterized the key concepts. These themes were based on the descriptive and interpretive codes, but involved a higher level of abstraction. The outcomes from the second cycle resulted in the answers to the research question under investigation.

#### **6.4.2 Quantitative data analysis**

For the quantitative data, statistical analysis of the repeated survey data was undertaken in order to determine students’ perceptions of their lecturer’s English. Because the repeated survey involved analyzing data on three themes at two points in time, a repeated measures ANOVA was employed. Moreover, ANOVA allowed me to examine the data for possible interactions between the themes across time. For example, if the values for one lecturer improved from first to last for one theme, such as language use, this factor would be seen at a glance. To observe the same interaction using a dozen or more t-tests would be much more difficult as well as additionally raise the probability of false positives<sup>18</sup>. The statistical analysis was performed with the assistance of Dr. Sebastian Pannasch, a visiting

---

<sup>18</sup> A false positive refers to a result that indicates a given condition to be present when it actually is not. In other words, the absent events yield positive test outcomes. Although mathematically a false positive ratio is equal to a type 1 error rate, it is viewed as a separate term. For further details, see *Statistical Analysis for Language Assessment*, Lyle Bachman 2004.

research psychologist at Aalto University. The data were analyzed using SPSS 20 (Chicago, IL). With respect to the presentation of the findings, the details are presented with the results in Chapter 9 in order to make it easier for the reader to follow the procedures while reading the results.

## 6.5 Validity and Reliability

The concepts of validity and reliability are well established in quantitative research, where they serve as the measure by which such research is judged. However, qualitative work belongs to a different paradigm that involves interpretation, an aspect that underscores the researcher as an instrument, which means that the qualitative findings will be subjective and contextual, to some extent. By nature, they are also inherently inconsistent (Dörnyei 2007: 57). In an attempt to overcome the differences between these two paradigms, Lincoln and Guba (1985) offered four criteria for qualitative research that define *trustworthiness*, which they view as comparable to the validity and reliability criteria in quantitative research: confirmability (or neutrality; parallel to ‘objectivity’), credibility (or truth value; parallel to the ‘internal validity’), dependability (or consistency; parallel to ‘reliability’), and transformability (or applicability of the findings to other contexts; parallel to ‘external validity’). While Lincoln and Guba advocated adopting these parallel qualitative criteria as corresponding to quantitative criteria, others argued against it on the premise that qualitative research (belonging to a completely different paradigm) must develop its own way of accounting for what is considered a valuable knowledge claim (for discussion, see Morrow 2005). At present, the emphasis for validity in interpretive, qualitative research is on trustworthiness (Cohen et al. 2007).

To establish trustworthiness in research, many researchers encourage a multi-methods approach that uses triangulation strategies, whether qualitative, quantitative, or mixed (e.g. Dörnyei 2007; Tashakkori and Teddlie 2003). The triangular strategy employed in the present study includes both method and data triangulation. These strategies were suitable for obtaining a more holistic view of the outcomes, reducing the chance of systematic bias, and increasing the credibility of interpretations.

While trustworthiness is the key concept applied to qualitative research, other efforts are still being made to identify validity concepts. For example, Maxwell (1992) introduced an influential taxonomy of validity for qualitative research, which consists of five concepts: descriptive validity (achieved through investigator triangulation), interpretive validity (accomplished through participant feedback), theoretical validity (corresponds to internal validity), generalizability (corresponds to external validity), and evaluative validity (similar to Bachman’s usefulness criteria). Of these concepts, the present study employed generalizability. In the qualitative account, the strategy used to examine generalizability was to incorporate the participants’ judgments of the generalizability of the targeted issue (Dörnyei 2007: 59), that is, whether the five selected CEFR scales were suitable for assessing their spoken professional English for an international context. It also employed interpretive validity of the outcomes, which was achieved by presenting and discussing the repeated survey findings with the respective lecturers. To ensure the

validity and reliability of the survey instrument, my doctoral supervisor and the students present at our post-graduate seminar meeting also reviewed the survey items. This step served as a checkpoint to ensure that the items were worded in a reader-friendly manner and that the questionnaire was suitable for the purpose. In addition, examining the lecturer's English from both the perspective of the lecturer and their students gives face validity to the findings.

While the taxonomies of Lincoln and Guba (1985) and of Maxwell (1992) are helpful in establishing principles for reliability and validity in qualitative research, the most important strategy for ensuring trustworthiness of a research project lies in the "image of the researcher as a scholar with principled standards and integrity" (Dörnyei 2007:59). In other words, the extent to which a study is trustworthy can only be determined by its readers, and their verdict will depend on the research design, the data gathered, and the skill of the researcher.



## 7 Analysis of language ideologies

This chapter focuses on analyzing the language ideologies that EMI lecturers uphold regarding their professional English. One aim of the analysis is to identify the ideologies that are foregrounded at any point. Another is to explore a connection between the language ideologies upheld and self-perceptions of professional English. For the investigation, the following hypothesis was formed: EMI-lecturers' language ideologies will guide their positive and negative self-perceptions of professional English. My focus in the analysis is on the beliefs and language ideologies that manifest from the data as related to the widespread language ideologies introduced in Chapter 3, and whether these beliefs inform the self-perceptions of professional English.

In the analysis, I present the language ideologies of my informants as being comprised of sets of beliefs about language use and social interaction. With this as the starting point, I have organized the chapter such that each section corresponds to one widespread language ideology in the theoretical framework.

### 7.1 One-nation-one-language ideologies

From the data analysis, one theme that manifested links to countries, and the related beliefs connect culture and/or language to a country. One ideology that emerged is a one-nation-one-culture ideology. This belief was present in discourse about 'being and sounding Finnish'<sup>19</sup> as part of the normal ethno-landscape in Finland whether speaking English or Finnish. This view connects sociolinguistic behavior to a particular context, in this case, to the local country. Another ideology that emerged is a one-nation-one-language ideology (see discussion in 3.5.1). This belief was present in discourse about speaking Finnish like a Finn in Finland and English like an ENL speaker in an ENL country. Although these two ideologies contrast to an extent, they also echo the local context (i.e. country) as being one norm of orientation – in one case through culture and in the other through language.

For instance, the accounts in 7.1-.2 present the ideology of one-nation-one-culture through maintaining sociolinguistic features that are seen as acceptable when speaking L2 English in the local country.

(7.1) <I> ok, so you said that you don't think you need to speak like a native speaker <OF ENGLISH>, why is that </I>

---

<sup>19</sup> From the data as a whole, references to 'being Finnish' seem to refer primarily to following Finnish politeness rules when using L2 English (applicable to Finns and non-Finns), and 'sounding Finnish' to having traces of a Finnish accent in L2 English (applicable to NSs of Finnish) and in L2 Finnish (applicable to NNSs of Finnish).

- <B> because i'm not a native speaker @i'm a finn@so, that's something i would like to be able to, to speak like a native speaker, but it's not necessary, i would be understood even if i don't speak like a native speaker, if i speak correct english </B>  
 ((...))
- <I> ok, and uh how about when you speak English, any concerns about others being able to recognize influence from your first language </I>
- <B> not much, at least not here, not in finland </B>
- <I> and how about otherwise, in other situations </I>
- <B> so, it depends it doesn't bother much but er I don't think it would be, would like to be recognized as a finn because of my way to speak english </B>
- (7.2) <I> how about, how would you describe your own accent </I>  
 <C> @@ i i guess @@ there must be a very clear finnish accent </C>  
 ((. . .))
- <I> are you concerned er about other people recognizing your accent as being a finnish </I> <C> [not in a] finnish environment, in <NAME OF CITY IN USA> yeah, I think it was to some extent, but not here </C>

Both accounts suggest that it is acceptable to 'sound Finnish' in Finland, but not necessarily outside the country. This view, however, is not the only one present. In both examples, the accounts imply a desire to speak English like a NS in ENL countries. These beliefs signal the presence of dualistic language ideologies, which are in conflict.

Moreover, the belief in being Finnish in Finland was expressed not only by Finns but also non-Finns, as the next accounts illustrate. In 7.3, the account suggests a SLU will attempt to uphold the politeness rules of the local country.

- (7.3) <I> how about when you feel the need to be polite, so what kind of rules do you follow </I>
- <A> i use a combination of three different rules, the first one is the one that is from my own own culture, and then because i also know very well about the french culture then i use i i combine those ones and then also because i have been for a long time in finland i i use those ones if i can </A>
- <I> even if there are not finns present </I>
- <A> ye yeah sometimes yeah </A>

While this account illustrates an effort to respect local politeness rules, it simultaneously demonstrates pluriculturalism. Rather than replacing previously learned rules with new ones, the account suggests that local rules have been added to the already existing repertoire. In this way, this example illustrates respect towards the local country and lends support to a one-nation-one-culture ideology.

As explained in the following account, acquiring a Finnish accent for speaking Finnish in Finland is also a target among non-Finns. This L2 speaker's view supports a one-nation-one-language ideology. A similar view, however, does not extend to SLU.

- (7.4) <I> ((...)) is it your target to pronounce english like a native speaker </I>  
<D> not really, my target is to pronounce finnish like a native speaker, english  
that maybe it's just good enough to survive in this environment </D>

What this account illustrates is dualistic views: one as a SLL of Finnish and the other as a SLU of English (see discussion on 'learner' vs 'user' in 4.2). Moreover, the latter view contrasts with SLA notions that support native-like targets (as discussed in 4.2). These conflicting views reflect a relationship between language beliefs and context of use.

The accounts in this section link language use to personal experiences in a Finnish context, where the belief in being Finnish in Finland is shared. With this belief as the guide, both positive and negative self-perceptions were expressed. Although it was acceptable to sound Finnish in Finland, it was not necessarily desired when speaking English. Moreover, being Finnish was limited to a Finnish context. Self-perceptions of sounding Finnish outside of Finland were not necessarily viewed as positive, and clearly seen as negative in ENL contexts. The beliefs upheld relate social and linguistic behavior to countries, and contrasting language ideologies are present. On the one hand, L1 language was related to the local country, as depicted by the discourse of speaking native-like Finnish in Finland and native-like English in an ENL country<sup>20</sup>. On the other hand, discourse related to speaking ELF did not support a one-nation-one-language ideology, but rather a one-nation-one-culture ideology as depicted by the desire to respect the local (i.e. Finnish) sociolinguistic behavior of the local country (i.e. Finland).

## 7.2 Standard language ideologies

Beliefs related to standard language ideologies (discussed in 3.5.2) emerged in discourse evolving around linguistic correctness and communication concerns. Thus, the analysis here focuses on correctness talk, which is what I take as expressing standard language ideologies. This section first introduces how correctness was talked about, after which it presents approaches to linguistic correctness when lecturing, and then concludes with how the talk about correctness relates to self-perceptions of language competence.

To gain understanding of perceptions of correctness, I asked my informants to define grammatical correctness and to explain their views on it. These topics were prompted in the CEFR interviews in relation to the grammatical accuracy scale. From the interview data, three primary views arose, one of which was prescriptive (as discussed in 3.5.2), another idealized (as discussed in 3.5.3), and the third performative (as discussed in 5.3.3). The prescriptive view upheld the belief that grammar books are the right source of correctness for linguistic form, as this example illustrates.

---

<sup>20</sup> Although language ideologies such as one-nation-one-language are widely upheld, it does not necessarily mean that a country has only one official language. Finland, for example, is bi-lingual with two official languages: Finnish and Swedish.

- (7.5) <I> and my first question is, what does grammatical accuracy mean to you </I>  
 <A> grammatical accuracy, er for me, so it would mean that er the nouns the  
 verbs and things like that are correctly put into a sentence and also  
 conjugated in the er the right tense </A>  
 <I> ok, and would it be according to some specific rules </I>  
 <A> i think yeah i think there are some standards which are the correct things to  
 do, so er *measuring according to grammar textbooks that is the right target*  
 </A>

The belief in grammar books as the source of correctness, however, was only one view held. As the next account illustrates, another model of correctness was the belief in the idealized NS.

- (7.6) <I> ok, how about then er when you're lecturing your target then in terms of  
 correct grammar would be </I>  
 <B> it would be *near native speaker* </B>

The belief in the idealized NS as a model of correctness reflects a learner-related language ideology, as discussed in 3.5.3.

The third view to correctness takes performative varieties as the point of reference, a view in opposition to prescriptivism and idealization. The performative view upholds the belief that correctness differs in spoken and written language (see 5.3.3). Explicit in this belief is the assumption that spoken language is imperfect because it contains dysfluencies not found in written language, as illustrated in this example.

- (7.7) <I> all right, how about, er what are your views on correctness in speaking in  
 english </I>  
 <D> well, *you should speak correct english but of course it's not the same  
 correctness than it is when writing, i mean you er in practice just, even a  
 native speaker will er deliver an even often slightly incompetent  
 performance when speaking english, i mean sentences are not finished and  
 words that are doubled up and so on and hesitations, that's okay* </D>  
 <I> ok, er all right, and what is your idea of correctness based on </I>  
 <D> erm well it was a long time ago to refer to what i learned at school, but  
*nowadays it's more what you see and hear in television and so on* </D>  
 <I> ok </I>  
 <D> *cause language is what people speak it's not what the books say it is* </D>

Associated to this view is the assumption that correctness in spoken language is not necessarily related to a set of rules from a grammar book, but rather to the views of the user as based on his/her experience. As accounts 7.7-8 indicate, this experience includes exposure to a wide range of accents via mass media and international contacts.

- (7.8) <I> and how about you, when you yourself judge grammatical accuracy, er yeah, how do you judge it yourself, or know when it's correct </I>  
 <D> that's a difficult question, i think *it's based on my internationalization of how one should speak english*, so it started in school and after that i used it in a professional setting, and it's how, it's how i remember it, it's not a set of rules at least, at one point in school there were rules that i remembered, but since then *i've been exposed to so much english, actually it's my inner ear that hears when somebody makes a grammatical error* </D>

What the accounts suggest is that the experience of using English has helped to develop an intuition about grammaticality that extends beyond the English learned at school. The account also suggests that the experience is international, indicating a wide array of exposure to different Englishes. Given that English is a global lingua franca, its large-scale use includes an extensive group of multilingual speakers who live outside an ENL context. This context suggests the absence of Standard English as the dominant language. It also implies that the “internationalized inner ear” of these SLUs of English will most likely include some linguistic variability not found in Standard English. In this way, the account in 7.8 hints at a non-standard norm as the measure of correctness. Such a norm, however, should not be understood as radically different from Standard English. In fact, a recent comparative study by Mauranen based on spoken academic data from ENL and ELF corpora shows ELF to be very similar to Standard English (2012: 247)<sup>21</sup>.

As this next example illustrates, the focus of communication extends beyond accuracy to appropriateness.

- (7.9) <I> er, how about what is your target, like for yourself, in terms of correct grammar </I>  
 <D> well, just correctness, speaking correctly, *formulating myself correctly, using that's mostly semantics, but using using appropriate terms appropriate words trying also to get shades of meaning* </D>

Although asked about grammatical correctness, this account shifts the focus of correctness from grammatical accuracy to lexical appropriateness. This shift implies a stronger focus on communicating meaning than on grammatical correctness, a view that aligns with the SLU paradigm (as discussed in 4.2).

All in all, what these accounts suggest is that there are different notions about linguistic correctness that are based on different points of reference, each of which provides a different standard against which to measure correctness: grammar books, the idealized NS, and the internalized rules accumulated through exposure to spoken language. With this as the point of departure, the following accounts illustrate beliefs about correctness in relation to communicating in English, where the actual spoken practices do not necessarily align with the prescriptive beliefs proclaimed (in 7.5-7.7). This difference becomes evident in the following examples.

---

<sup>21</sup> Note that this study did not include phonology in the investigation.

In 7.10, a prescriptive belief is explicitly present. The account associates incorrectness with ‘bad’ English, which is perceived as irritating and unprofessional.

- (7.10) <I> and er, does speaking correctly have anything to do with sounding more professional, or is it just a desire to speak correctly </I>  
<B> yes, i think it does, and i want to speak correctly, *i don’t like it when i hear bad english, it irritates me* </B>

A less strict view is present in the following accounts, 7.11-.12, which emphasize the importance of correctness in spoken language. In 7.11, the belief in speaking correctly reflects a prescriptive, correctness-orientation to language use. Explicit in this belief, however, is lenience as signaled by the allowance of incorrect prepositions and articles, for example.

- (7.11) <I> well, more about your view on it, what do you think about speaking correctly</I>  
<B> *i think i should speak as correctly as possible, but i can also be understood if i don’t especially*, well i may forget an article or forget a preposition but that, well it’s not completely wrong, i try to use the correct ones correct articles but sometimes i forget </B>

What this account signals is a contrast between a correctness goal and comprehensibility standard, where grammatical imperfections are seen as unimportant to the main goal. Not only does this view contrast with the prescriptive views proclaimed, it also supports an SLU paradigm (see 4.2).

Unlike the previous account, the following accounts view correctness as having an efficiency function in spoken communication. More specifically, the account in 7.12 suggests the motivation for speaking with correct grammar as one way of compensating for variation in pronunciation, which can be a problem – especially in ELF interaction where variability is high. In this way, the account suggests intelligibility to be a function of correctness.

- (7.12) <I> and how about, do you need help with grammar or vocabulary </I>  
<A> my vocabulary is good except for the pronunciation of some words, i also know i make some grammar mistakes, if you could point out the systematic mistakes, i could learn those . *speaking with correct grammar can make the message clearer especially since pronunciation can be a problem, so it’s important to try to be as correct as possible* </A>

The belief in communicating efficiently is also present in the next example, where self-correction in lectures is limited to correcting meaning, not linguistic form.

- (7.13) <I> <STIMULATED RECALL> so then, i was wondering er whether you monitor your language structures when you’re lecturing</I>

<A> i think i monitor it all the time, so so many times when even if i'm very much concentrated on on the subject then *after after i said the sentence i notice there was an error in that sentence but i never come back of course to that one to correct it, i just continue and i notice that in the previous sentence i said something wrong* </A>

<I> ok </I>

<A> it could be it could be the grammar it could be the vocabulary it could be also in the subject itself </A>

<I> ok </I>

<A> *ok if it's related to understanding the subject then i i'll correct it of course* </A>

<I> right </I>

<A> *but if it's only the english i i don't go back to it* </A>

The account in 7.13 suggests an awareness of making linguistic mistakes while lecturing, which are ignored unless perceived as important to comprehending the subject matter itself. This behavior suggests that self-correcting other aspects is viewed as unnecessary. Thus, minimizing self-correction of linguistic mistakes appears to be a strategy to support communicative efficiency, and contrasts with the ideology of prescriptivism that was earlier proclaimed.

Communicating efficiently is also implicit in 7.14. This account reflects a belief in clarity over accuracy, a view that directly opposes prescriptivism. This view, however, does not annihilate correctness as a factor in communicating efficiently. Rather, it assigns less importance to correctness.

(7.14) <I> all right and er what are your views on correctness when speaking English </I>

<C> i try to be correct to some extent, but *i'm not so much correct, i'm not so concerned about that any more*@i used to be more, so </C>  
(...)

<I> right, ok, er so *what is the main point for you*</I>

<C> *clarity* </C>

<I> clarity, ok, yeah </I>

<C> clarity, how how clear i can be *how clearly i can express my ideas* </C>

<I> ok </I> <C> *[not the] accuracy aspect* </C>

<I> alright, let's look at </I> <C> *[i can really] imagine that i could be totally accurate but er unclear* </C>

The account suggests that having 'perfect' language does not ensure clarity of expression. Hence, this account contrasts sharply with the account in 7.12, where intelligibility (at least overlapping with clarity) was perceived as following from accuracy.

Thus, although three different measures of correctness were proclaimed (7.5-7.7), the accounts suggest a belief in communicating efficiently taking precedence in spoken interaction. In other words, it appears that adherence to prescriptivism serves different

kinds of communicative purposes. In this way, the accounts indicate different relationships between successful communication and accuracy. The picture that emerges is complex. It includes both strong views on the virtues of correctness and more practically oriented views of other matters like ‘clarity’ and ‘lexical appropriateness’ overriding it. The latter views suggest prescriptivism to be less important than communicative efficiency in ELF communication. Nor do the views correspond to an idealized NS model of correctness. They do, however, seem to support adhering to a performative view that has an orientation to communicating efficiently without ‘perfect’ English. This orientation illustrates the belief in alternative norms for ELF.

The final part of the analysis investigates whether prescriptivism guides self-perceptions of professional English. The accounts suggest prescriptive beliefs to be upheld to different extents. Nevertheless, with prescriptive beliefs as the guide, users evaluated their English negatively, perceiving it as having fallen short of the standard language target. When adhering strictly to prescriptivism, there was less tolerance for linguistic inaccuracies. With this strict view as the guide, departing from ‘perfect’ English was viewed critically and seen as ‘bad’ English. Alternatively, when guided by a lenient view to prescriptivism, minor differences in usage of articles or prepositions, for instance, were perceived as acceptable. While prescriptivism was present in the beliefs, it was not the only ideology. A belief in communicating efficiently also emerged. With this ideology as their guide, users assessed their English more positively. For instance, the account in 7.15 indicates satisfaction with speaking correctly enough, which suggests that correctness is not the main focus.

- (7.15) <I> uh, how do you know if it's accurate or not </I>  
 <C> @for the spoken language, or </C>  
 <I> yeah </I>  
 <C> i haven't thought so much about that, ok, *sometimes i i'm aware of the problems, ok, for me, it's enough if i'm satisfied with enough that there's not too many errors, so, level of let's say, yeah, now (xx) i i am not so, er think so much about the grammatical correctness er when i'm writing er technical technical, then it's different </C>*

The following self-evaluation also suggests that ‘perfect’ English is not expected by the (primarily) NNS students, and even views imperfect English to be a merit.

- (7.16) <I> ok, i was wondering about er what kind of image you have of yourself as a lecturer </I>  
 <B> *i think it's good that students know that i'm not perfect, so, i could be a good example i can be understood even if i'm not correct, i am a perfectionist so i try to be good in some things but not in all@ </B>*

The account suggests that demonstrating the ability to deliver a lecture with imperfect English is a good model for EMI students, as it shows that ‘perfect’ language is not necessary for comprehensibility.



Overall, the picture that emerges is one where the primary focus is on communication, with linguistic accuracy assuming a secondary role. The accounts also illustrate awareness of English being imperfect, yet they indicate confidence in the ability to lecture in English.

### 7.3 NS language ideologies

In this section, I discuss language ideologies related to L2 proficiency targets, which evolve around beliefs connected to the idealized NS as a model. This idealized view is present in discourse that talks about NSs as though they are all the same. The views, however, are not straightforward and indicate a love-hate relationship with the idealized NS. While one view presents NS targets as desirable and appreciated, another disparages the NS and views NS targets as impractical.

Discourse related to the NS emerged primarily from three different interview prompts, which inquired about the following: 1) L2 proficiency targets, 2) the CEFR scales, and 3) experiences in using English in professional contexts. While some informants clearly specified their L2 targets as near NS, others targeted ‘good enough’. Moreover, as the accounts will illustrate, the discourse related to NS accents ranged from envy to scorn.

#### 7.3.1 A hierarchy of accents

The contradictory beliefs in NS targets were most prevalent in discourse on accents, where the accounts suggest a hierarchy of accents. This hierarchy places ‘neutral’ accents at the top and ‘strong’ or ‘incomprehensible’ at the bottom. The hierarchy includes NS and NNS accents. A related belief upholds some kind of neutral accent as one that is clear and intelligible. Another belief upholds accents that sound ‘natural’ and ‘nice’ as desirable, and those farthest from this ideal as undesirable. As the following account illustrates, an undesirable NNS accent is one that lacks natural flow.

- (7.17) <I> all right, erm, do you like your own english accent </I>  
<C> @@if i listen to it probably not, most likely@@ </C>  
<I> all right, why do you think you wouldn't like it </I>  
<C> er because, anyway *i somehow appreciate even when non-natives are speaking english some kind of natural flow or something that i, that i think i'm missing to some extent*@@ </C>

In the example, it appears that natural ‘flow’ means ‘fluency’, which is seen as a desirable feature. The account also suggests that NSs possess it as well as some NNSs. Appreciation of a clear NNS accent is also present in the next account.

- (7.18) <I> ok, so what do you think is important for clarity </I>

<B> *just this neutral way of speaking*, i also envied this <NAME OF SCHOOL> retired english teacher <AT EVENT WHERE I WAS ALSO PRESENT> her accent is *really nice* </B>

This example describes perceived clarity as including some sort of neutrality. It also mentions an “envied” accent, belonging to that of a NNS teacher of English. Although the account describes this accent as “really nice”, I was also present at this event and would add that this accent had clear pronunciation with a mild British intonation<sup>22</sup>. Given this interpretation, it appears that ‘neutral’ is the key word for clarity and that native-likeness is enviable but not necessary. This next account also emphasizes neutrality over native-likeness.

(7.19) <I> all right, and to be intelligible in your lectures, do you feel that you need to speak like a native speaker </I>  
<B> *no, no, it might even be better to be accent-free or ordinary, this kind of not native but clear* </B>

What seems to emerge from these accounts is that native-likeness is less important than some sort of neutrality, which contributes to perceived clarity and perceived intelligibility.

Moreover, accents that depart from these views lead to negative perceptions as the following account illustrates.

(7.20) <I> all right, erm i’ll go back to this accent question, to be successful in lecturing what kind of accent is needed </I>  
<B> *something that is not awful* </B>  
<I> what would you describe as awful </I>  
<B> a week ago i was in this lecture and *the speaker had a very strong finnish accent* </B>  
<I> so a strong accent </I>  
<B> *i had to leave because i felt i didn’t like it, i understood him, everything, but i’m also very strict to others* </B>

In 7.20, the account illustrates negative feelings caused by a strong foreign (i.e. Finnish) accent. It implies that strong NNS accents are bad and thus undesirable. Not only were strong foreign accents perceived as awful, but also incomprehensible NS accents, as the account in 7.21 illustrates.

(7.21) <I> all right, er, how would you describe the accent of a native speaker of english </I>

---

<sup>22</sup> *Although to my ear this accent sounded like a fairly standard pronunciation with some kind of British intonation, I do not know British accents well enough to identify whether it could qualify as RP.*

<A> *i guess when they are speaking between them it's good but when they are speaking to non-native speakers i think it's, sometimes it's awful, they don't care that the others are non-native speakers so they, erm almost every time when i go to conferences the chairman of the session say please remember that we are not native speakers and speak slowly and understandably and this holds especially for native speakers, so so usually when when you are at a conference you can understand everybody except Americans, and then you have to concentrate very much before you can start to understand and some british also do it, not all of them, when when they speak the professors speak in conferences they speak very good, very good English, but some of them they they speak the probably the spoken language in england and then then it's very hard to follow them* </A>

In this example, it seems that ENL speakers are viewed as having very good English that is not very intelligible to NNSs. The account also suggests the lack of comprehensibility to stem from ENL speakers talking too fast, in other words, not bothering to accommodate NNS listeners. This behavior implies lack of experience in using English with NNSs. The account also hints at strong (local) NS accents as being problematic.

What seems to emerge from the accounts is that 'good language' ideals do not equate with what is perceived as clear, intelligible communication. This difference contrasts sharply in the accounts. For example, the strong NNS accent in 7.20 was perceived as comprehensible although not liked (by a NNS who shared the same L1). In contrast, the account in 7.21 perceived a NS accent (considered to be good English) to be incomprehensible. Unlike these two views, the native-like accent in 7.18 was perceived as desirable, but neutrality was deemed more important for clarity (7.18-19). These different views suggest that nativeness is not the key feature that guarantees clarity or comprehensibility in NS-NNS interaction. Rather, some sort of neutrality is placed at the top of the hierarchy. Ideal neutrality, however, may not be easy to define. According to the accounts, it seems to include acceptance of variation in accents, good pacing (e.g. not too fast), and a comprehensibility goal (e.g. able to produce most of the sounds of English, as discussed in 7.3.2, and to understand them).

On self-perceptions of professional English, a deficit view is also applied to undesirable aspects of NNS accents, as this account illustrates.

- (7.22) <I> ok, all right, erm how would you describe the accent of a native speaker </I>  
 <B> so, for example *finnish is quite dull, so you should be able to to use stress more more correctly in the places on the words, i think it's not very difficult for me because i think i learned that, sometimes when i hear other finns speak and they can't i don't know some words finns can't pronounce some words they always say them in the wrong way@@* </B>  
 <I> ok, erm and so it bothers you </I>  
 <B> it bothers me a lot </B>  
 <I> how about your own accent, how would you describe it </I>

<B> *maybe dull finnish@@finnish English@@* </B>

<I> *dull finnish english, meaning* </I>

<B> *so, my intonation is so dull so it's very hard for the listeners to, and i speak a lot too in in the early mornings to to stay awake listening to my english* </B>

The deficit views in 7.22 (as well as 7.20) reflect findings from previous studies where L2 speakers with the same mother tongue will be the most critical of each other, especially on identifiable mother tongue features stigmatized as common mistakes through traditional language teaching (e.g. Lavelle 2008). Consequently, many L2 speakers of English may see themselves and their fellow countrymen as failed learners of English, in accordance with NS language ideology, an ideology to which all SLLs are exposed at school.

Conversely, the next example illustrates a non-hierarchical view to a NNS accent. In 7.23, a NNS accent not viewed as strong (cf. 7.19) is seen positively. This belief illustrates a differential view to NNS accents as opposed to a deficit view, as this account illustrates.

(7.23) <D> <STIMULATED RECALL> *i see i have a noticeable dutch accent to my own ear, i'm not sure that others hear that, so, it's not so dangerous* </D>  
<I> *you said, 'it's not so dangerous'* </I>  
<D> *no, not so dangerous, but if it's strong it may get in the way* </D>

This account suggests the presence of a common-sense view, where the speaker perceives his accent to be neutral to others, as signaled by the comment about others not noticing his Dutch accent. This view is inherent in the belief that 'you have an accent but I don't'. This common-sense view is also present in the ideal NNS accent as being 'ordinary' or 'accent-free' in 7.19.

What these accounts illustrate is a difference in self-perceptions that depend on language ideologies. Guided by NS language ideologies, accents with native-like features were viewed positively (e.g. 7.17-18) while those that differed greatly were viewed negatively (e.g. 7.20 and 7.22). Conversely, when guided by a common-sense view, NNS accents were perceived positively (e.g. 7.23). This latter view is also reminiscent of the non-hierarchical notion present in English-as-a-global-language ideologies (see 3.5.4 for a description).

### 7.3.2 NS – idealized vs real

In discourse on pronunciation, more contradictory views manifested. One viewpoint sees NS targets as a good model, a belief signaling NS language ideologies. A contrasting view, however, regards attaining native-like pronunciation as impractical for ELF communication. Moreover, the belief in ELF users being able to communicate as well as NSs, if not better, is also present. This belief suggests the absence of a hierarchical view

towards ELF communication, which signals the presence of English-as-a- global-language ideology, discussed in 3.5.4.

The accounts in 7.23-7.25 suggest a belief in NS targets being irrelevant and useless for ELF communication. This view is present in 7.24, where the account suggests that trying to attain native-likeness is not worth the enormous effort. Rather, what is believed to be important is to speak in a way that enhances communication.

- (7.24) <I> ok, all right, so, *you said it's not your goal to try to speak like a native speaker* </I>  
<A> erm </A>  
<I> and i was wondering *why it isn't your goal* </I>  
<A> er, i think it's just a, it's just on one hand *it's quite difficult, you have to be rehearsing all the time and listening to native speakers all the time* and so on and on the other hand *it's not that useful, so you lose a lot of time learning something which is which is basically not useful because because you can make yourself understood without doing that@@i think the main point in in language is to enhance communication between people, and now if you can communicate with with less effort as good as as well as a native speaker or even sometimes better than a native speaker then then, why why should you try to imitate a native speaker* </A>

Like 7.24, the accounts in 7.25-.26 emphasize a belief in the importance of communication. They, however, contain conflicting views about the NS. In 7.25, the account implies leniency towards accent variability, and then targets error-free pronunciation that is 'more' native-like than that present in this speaker's self-perception. It also faults NSs for excessive speed.

- (7.25) <I> what do you think about correct language, is it important to speak correctly </I>  
<A> no, i think that the most important point is to be able to communicate your ideas and to know that you have been understood, perfect grammar is not important but pronunciation can be a problem, *you know there are many kinds of english in the world today and many of these local accents as well like they have in asia, this is why i want to improve my pronunciation to be more native-like*, if you can tell me my systematic errors i know that i can improve them, *but i don't need to speak like a native speaker*, you know these *americans at conferences*, you can't understand anything they say when *they speak so fast, you ask them to slow down and they do for a moment and then they are right back at high speed again, they don't seem to understand why non-native speakers don't understand them, they have only one way of speaking* </A>

Explicit in this account is the need for accommodation in international communication as well as the belief in speakers being able to adjust their communication in NS-NNS

interaction in international contexts. The account also suggests a desire to be more native-like in order to be able to cope with the different varieties of English. In other words, the more native-like the pronunciation, the better it fits all purposes as this would help to reduce variation in pronunciation. Similar to 7.24, this account also views native-likeness as good but not necessarily worth the huge investment. Conflicting views between idealized and real L2 communication targets are also present in 7.26.

- (7.26) <I> do you think that measuring your own accent against a native speaker is relevant for your use of english at work </I>  
<A> *no, no, it's not relevant, although it could be, it could be good, but it's not relevant, yeah it could be a good reference where you can get most of the right pronunciations, but you should put it in its framework somehow so that you don't exaggerate it, i don't have to compare how how i am speaking with a native, they're they're speaking a different way anyway* </A>

What seems to emerge from these accounts is a clear distinction in the views toward pronunciation as related to a NS model and the NS as an actual participant in international communication. Implicit in the NS-as-a-good-model belief seems to be the assumption that targeting a native-like pronunciation will help to reduce variation in ELF accents. Another assumption appears to be that achieving this target does not entail speaking in the same manner as a NS, which is viewed as inefficient in ELF communication due to the lack of accommodation. These assumptions illustrate a contradiction between an idealized view of pronunciation and the actual experience of communicating with NSs.

On self-perceptions, the accounts suggest that taking NS language ideology as the guide for pronunciation leads to negative perceptions when falling short of idealized NS targets. This view is signaled by the desire to improve in order to become more native-like. Conversely, the opposite is implied in the accounts, that is, pronunciation that closely approximates native-like targets results in positive self-perceptions. These views were present even though NS targets were perceived as impractical and useless for ELF communication.

### **7.3.3 Different sociolinguistic appropriateness for different social contexts**

What the next accounts illustrate is different sociolinguistic appropriateness for different social contexts. They also show that a speaker may be subject to pressure to conform to a particular standard in a given context, where failing to meet the standard results in feelings of alienation. For contexts to which this applies, I refer to the sociolinguistic appropriateness as externally imposed and as representing a macro view (i.e. at the level of group). In contrast to this is a micro view at the level of individual. This view of self can extend beyond an actual context to one that is self-imposed. Either way, self-imposed (micro) or externally imposed (macro), context dependency is central to views towards sociolinguistic appropriateness.

For instance, the account in 7.27 suggests that using English in an international context does not arouse feelings of needing to conform to ENL norms. Conversely, this feeling does exist in ENL contexts. These views illustrate a belief in sociolinguistic appropriateness being tied to context at a macro level.

- (7.27) <I> ok, so, *here in this environment, what is the difference for you, why here is it ok* </I>  
<C> *there are so many different kinds of accents, so, so it's er, i think i spoke good, well enough, so, so* </C>  
<I> ok, so *this is more international with lots of different accents* </I> <C> [but then i am going to] ok, not, i'm not so much worried anymore, but definitely i i, this is the most, *i i think the most difficult place is England, definitely, definitely much more there, i i think there, there is some feeling that i should speak more clearly and some oxford movies and@so, if you are not speaking well enough, you cannot be accepted, actually, in in the us <USA>, i, even there, there are so, so many non-native speakers <SIGHS> not, not so big problem@* </C>

According to the account in 7.27, a wide variety of accents are both present and accepted in an international context. In the account, acceptance is indicated by the positive view this speaker has of his NNS accent in this context, described as “good enough”. However, in an ENL context, such as England or the USA, the account suggests sounding non-native to be unacceptable. His experience suggests feelings of being stigmatized by a deviant accent and consequently unaccepted by the group.

This next example suggests the image of self as a speaker can change when the context changes. Before a primarily NNS audience, the account portrays self-confidence and positive self-perceptions in using professional English. However, changing the context to a primarily NS audience hints at less confidence.

- (7.28) <I> yeah, ok, I i guess i'm thinking now about something else, that er with *the idea of speaking correctly erm i was wondering if it's also related to sounding professional* </I>  
<D> you mean the image aspect </D>  
<I> yeah </I>  
<D> erm *actually before this audience i don't think that matters very much, they know i'm not a native speaker and anyways i speak better than them so so i don't think it really matters much, but perhaps in an environment where i would address native speakers it would be different* </D>

Similarly, the following account suggests that self-perceptions of professional English change depending on the point of comparison, where context is implied. The account in 7.29 suggests that taking other NNSs as the point of comparison against which to measure language competence leads to positive self-perceptions, whereas the accounts in 7.20 and 7.22 indicate taking the NS as the point of comparison leads to negative self-perceptions.

- (7.29) <I> all right, er do you like your own English </I>  
 <B> yes@@ </B>  
 <I> @@ok@@ er </I>  
 <B> so, it depends with whom you compare yourself with, *if i compare myself with ordinary finnish people then i am better or i think i should be pleased with myself* </B>

In this account, context dependency was present in the chosen measure, where choosing an “ordinary” Finn as the reference provides an alternative context for evaluating English. This account illustrates a self-imposed view, one that may be present (as rooted in past experience) and irrespective of the actual context of usage. This view represents a micro view (of self) as opposed to a macro view of self as part of the actual local context.

On self-perceptions, the views changed according to context as well as to the chosen point of comparison. At the macro level, English that deviated from ENL norms felt stigmatizing. Conversely, in non-ENL contexts, English that departed from ENL norms was viewed positively, as illustrated by the acceptance of a NNS (Finnish) accent as being ‘good enough’ and by the positive image of self before a NNS audience. At a micro level, context-dependency was present in the chosen point of reference. In other words, the point of comparison for evaluating English could be determined by a self-imposed context (rooted in past experience) that differs from the actual context (at the group level). For instance, a speaker could choose the NS or the average Finn as the point of comparison in any context where they use English. This chosen measure (that is context-dependent) leads to different self-perceptions. As illustrated in the examples, the NS as the measure led to negative self-perceptions of English whereas the average Finn led to positive self-perceptions of English.

## 7.4 English-as-the-global-language ideologies

One of the ideological assumptions present is that using English in international programs is a natural tendency. When asked about teaching through English, the accounts indicated acceptance of this choice, as illustrated in this example.

- (7.30) <I> what do you think about this idea of having so many international programs in english </I>  
 <B> i think it’s a natural tendency </B>  
 <I> a natural tendency </I>  
 <B> yes because of aalto university </B>  
 <I> ok </I>  
 <B> i don’t have anything against it, because i think it’s something they want in the university that’s why i’m coming to this mentoring program it’s planned and i accept it </B>



Furthermore, their acceptance of using English as the language of instruction in international programs also reflects a belief in using English in ‘the world at large’. This belief entails using English primarily with NNSs as a contact language (i.e. ELF), as the account in 7.30 indicates.

- (7.31) <I> <SUMMARIZING> ok, so you can be even clearer than a native speaker </I>  
<A> *yeah, so maybe we have to put it in its context that's er that's erm i don't know, there's tens or hundreds of millions of native speakers, i don't know how much but there's almost more than 1 billion of non-native speakers which need to communicate in in English, so when you are communicating you should understand that most of the time you are communicating with non-native speakers and very few times with native speakers </A>*

The account suggests the use of English to be a common-sense choice among NNSs who share it as a language for communication in contexts where NSs are seldom present. In such contexts, beliefs in being understood as well as in maintaining the L1 background also exist, as this account illustrates.

- (7.32) <I> ok, and erm is this because as you mentioned a few minutes ago that primarily you're speaking with other non-native speakers </I>  
<A> *yes, that that's one thing and the other, yeah mainly it's that i need to communicate with non-native speakers but sometimes i need to communicate with native speakers, i don't have to hide that that i'm not a native speaker, so i can speak with my er background provided that i make myself understood and don't make a lot of mistakes that they wouldn't understand </A>*

Present in the account is the belief in a comprehensibility goal when communicating in English. Another belief present is related to maintaining the L1 sociocultural background. The latter belief underscores diversity as a desired characteristic among lingua franca speakers, where there is no motivation to acculturate to monolingual NS norms. Rather, it suggests an ideology of norming towards multiculturalism.

Norming towards a lingua franca is also present in the belief in ELF users communicating with non-standard L2 English. Underlying this belief is the assumption that non-nativeness is accepted and even expected – not only by EMI lecturers but also their students. This assumption directly relates to shared non-nativeness among ELF users, as the account in 7.28 illustrates. Shared non-nativeness also points toward a differential view to English as opposed to a deficit view. In this non-hierarchical view, shared non-nativeness appears to be a binding factor.

The belief in shared non-nativeness is also present in the notion of utilizing plurilingua-cultural resources. Illustrating this point, the following account describes employing plurilingual strategies as a way to maintain fluency when lecturing.

- (7.33) <A> <STIMULATED RECALL>((...)) ok, for example, *for me the french grammar comes in from time to time, so, when when i don't know how something in english grammar, how it should be done, then i try i try to somehow reflect from french grammar*, and i try to put it there </A>  
 <I> yeah </I>  
 <A> sometimes it's ok, sometimes it doesn't work </A>  
 <I> yeah </I>  
 <A> because, you you get, *you cannot just stop and do nothing*, you have to do something </A>

What the accounts suggest is that drawing on L1 resources is a natural choice for communicative efficiency in NNS-NNS interaction. Moreover, the choice reflects the sociocultural experience of the user as well as a positive effort towards enhancing communication among speakers with diverse L1 backgrounds.

In short, the English-as-a-global-language ideologies contrast with language ideologies that support monolingualism, such as standard language, NS language ideologies, and one-nation-one-language. When self-assessing professional English in relation to English-as-a-global-language ideologies, English language competence was perceived positively, viewed as 'good enough' to be understood and as superior to that of NNS students.

## 7.5 Conclusion

In this chapter, my aim was to investigate the beliefs and language ideologies about English associated with the language competence of EMI lecturers. The basis for the analysis was the theoretical framework in Chapter 3. For the investigation, it was hypothesized that EMI-lecturers' language ideologies will guide their positive and negative self-perceptions of professional English. Four widespread language ideologies were discerned as well as their related beliefs: one-nation-one-language ideologies, standard language ideologies, NS language ideologies, and English-as-a-global-language ideologies. Although several widespread monolingualizing language ideologies were present in the data, the analysis does not suggest such ideologies as dominant in all contexts.

An important conclusion from the analysis is the clear evidence of a connection between language ideologies and context. Context-dependency was found to be central to the beliefs espoused about language usage in relation to social behavior, which was grounded in the experiences of my informants. One context was at the level of nation-state, and the sociolinguistic beliefs related to it. In this context, two language ideologies emerged: a one-nation-one-language and a one-nation-one-culture ideology. These ideologies included beliefs about non-Finns' desire to speak native-like Finnish and non-ENL speakers to speak native-like English (a sign of social integration to the host country), as well as beliefs about it being acceptable to act and sound Finnish when speaking English in Finland, but not outside of it. These beliefs connect the construction of social agency to a national context. Similarly, standard language and NS ideologies

were upheld for ENL contexts, where the feeling of needing to conform to NS norms was present. This feeling, however, was not present in non-ENL contexts, as shown in the one-nation-one-culture ideology. These examples illustrate the close relationship between context and agency, where the language ideologies upheld appear to be connected to whether the context arouses feelings of needing to conform to a particular norm (i.e. the macro level). I have referred to this situation as one that is externally imposed. Conversely, an individual could also uphold certain language ideologies through a self-imposed context. This latter view included a self-imposed point of comparison that provided a view of self in a chosen context as rooted in previous experience. In this chosen context, the point of comparison could be the NS or the 'ordinary' Finn, for example. In the case of self-imposed contexts, the macro context of self as part of the actual local context was irrelevant. The analysis also presents evidence of certain language ideologies being upheld even when individuals' interests may not be served by it. In an international context, the analysis indicates norming towards plurilingua-culturalism. In short, whether context was internally or externally imposed, language beliefs changed when the context changed.

Self-perceptions also changed when the context changed. For ENL contexts, English was evaluated negatively when it did not meet NS norms. Alternatively, in non-ENL contexts, English was evaluated positively against other NNSs, with self-perceptions suggesting the professional English to be 'good enough' to be understood and to be better than that of other NNSs. Similarly, the same phenomenon applied to changing the point of comparison from an idealized NS to an 'ordinary' NNS, for example. Likewise, EMI lecturers evaluated their English to be better than that of their NNS students, but seemed less confident about their L2 English before a primarily NS audience.

Another important conclusion from the analysis is the evidence of multiplicity of views illustrating the presence of different language ideologies, some of which were conflicting. For example, the views to standard language ideologies were contradictory. Although three different measures of correctness were proclaimed, the accounts about language use presented a belief in communicating efficiently as taking precedence over linguistic correctness. This belief included being understood as the main criterion over the prescriptive views that were also declared. Conflicting views to NS language ideologies also had a strong presence. Beliefs included in NS ideologies positioned native-likeness as appreciated, but not necessarily worth the huge investment. The discourse on accents also suggested a language hierarchy into which individuals could be indexed according to how closely they approximated a so-called 'neutral' accent. Positioned at the bottom of the hierarchy were accents viewed as 'strong' or 'incomprehensible'. Such a hierarchy constructs boundaries between individuals and attributes values to them, valorizing the desired accents and demonizing others. In sharp contrast to this language ideology is the English-as-a-global-language ideology, which is non-hierarchical. Beliefs included in this ideology are non-nativeness and plurilingua-culturalism. As far as the language ideologies expressed, my analysis shows dualistic views to be present, that is, NS language ideologies and English-as-a-global language, as well as standard language and English-as-a-global language. However, as a contender for international communication, the only serious candidate appears to be English-as-a-global language, as the related ideologies

clearly link it to ‘the world at large’ (also a contextual notion). However, this is not the full picture as the idealized NS is viewed as a good model for pronunciation, where targeting this standard is perceived as a means of easing variety in ELF accents. This idealized view, however, does not appear to be connected to a target culture, which seems to suggest a model for pronunciation that is devoid of native-likeness as the standard measure. This idea is reminiscent of Jenkins’ lingua franca core, for example. In addition, the prescriptive view to speaking inherent in standard language ideology was perceived as a strategy for enhancing communication, due to high variability in ELF.

The analysis also examined whether there is a causal link between language ideologies and self-perceptions of language use. A hypothesis was generated for this connection based on the following. At the core of standard language ideologies with prescriptivism intertwined are the rules to which speakers must adhere in order to produce language deemed as ‘correct’. Closely related to these ideologies are NS language ideologies. For NNSs, this means that language competence deemed as correct is ‘native-like’ and that which is not is ‘deficient’. Due to this edict, it was hypothesized that EMI lecturers would have positive or negative perceptions of their professional English depending on how closely they perceived it to approximate that which they considered to be the standard of (NS) correctness.

For H1 support was found. The analysis showed that self-perceptions of professional English changed, depending on which language ideologies guided the evaluations. Guided by one-language-one-nation ideologies, the accounts illustrate the desire to speak native-like language as linked to a nation-state. This view was present in accounts that perceived speaking native-like Finnish in Finland and native-like English in ENL countries as positive. Moreover, when standard language ideologies were the guide, lecturers viewed their professional English negatively, since it did not meet the standard. Two views to prescriptivism were present: strict and lenient. Adhering strictly to prescriptivism led to critical views of NNS English being imperfect, whereas adhering less strictly suggested a higher tolerance for imperfection. The ideology of standard language, however, was not necessarily upheld as reflected in a belief in communicating efficiently taking precedence over correctness. Taking this belief as the guide, self-perceptions of professional English were more positive, with non-standard English seen as satisfactory and even viewed as a good model for EMI students, as it shows the ability to communicate at a professional level without having ‘perfect’ English. With NS language ideology as the guide, users evaluated English positively when it met the idealized NS target. Conversely, they evaluated it negatively when it fell short of the target, even when NS targets were considered to be impractical, useless, and irrelevant for ELF communication. In short, self-perceptions of professional English tended to be negative when guided by language ideologies rooted in prescriptive, normative beliefs and positive when guided by language ideologies based on variation in language use.

## 8 Analysis of self-assessments in relation to five CEFR scales

My aim in this chapter is to investigate the set of five CEFR scales used in a performance-based pilot-certification test to assess the professional English of EMI lecturers. (For details on the assessment, see Chapter 2.) For this test, language competence was defined as C1<sup>23</sup> on the following CEFR scales: fluency, coherence and cohesion, vocabulary range, grammatical accuracy, and phonological control. These scales are presented in Chapter 5, along with an inspection of the ideological basis upon which they were constructed. This critical examination forms the theoretical framework for the data analysis in the present chapter.

My research question for the second part of this study investigates the CEFR on the basis of empirical analysis. More specifically, it examines the extent to which the five aforementioned CEFR scales are suitable as a measure of language for the intended purpose and situation. It was hypothesized that the CEFR scales will not tap into the skills relevant to professional English for the purpose of lecturing in an international context. The five scales are investigated through EMI-lecturer self-assessments of their professional and non-professional English and a semi-structured interview. The objective of the interview was to gain insight into how suitable lecturers found the scales as an instrument of self-assessment for evaluating their professional English.

First, this chapter briefly presents EMI-lecturer self-assessments, after which it analyzes data from the semi-structured interviews. Following the analysis, the chapter presents CEFR limitations for ELF and LSP usage, and then ends with a summary.

### 8.1 CEFR self-assessed English

The aim of this section is to establish lecturers' perceived differences in their professional and non-professional English as determined by their self-assessments. Table 8-1 shows lecturers' self-assessments on each of the 6-point CEFR descriptor scales, ranging from A1 (lowest) to C2 (highest). In the left-most column, the table shows each lecturer and his/her L1. The remaining columns show the names of the descriptor scales, with categories just below each for professional (work) and non-professional (gen) purposes containing the ratings.

---

<sup>23</sup> An exception to this was the level for phonological control: set at B2. Nevertheless, the overall pass mark was C1.

**Table 8-1. Self-assessed English on five CEFR scales for professional and non-professional purposes.**

	Cohesion & Coherence		Spoken Fluency		Grammatical Accuracy		Phonological Control		Vocabulary Range	
	Gen	Work	Gen	Work	Gen	Work	Gen	Work	Gen	Work
Lecturer/L1	Gen	Work	Gen	Work	Gen	Work	Gen	Work	Gen	Work
A/Arabic & French	C1	C2	B2	C1	B2	B2	A2	A2	B2	C2
B /Finnish	B2	C1	B1	B2	B1	B1	B2	B2	B2	C1
C/Finnish	B1	B1	B1	B2	B1	B1	B1	B1	B1	B2
D/Dutch	B1	B2	B1	B2	A2	B1	A2	B1	B1	B2
E/German	C1	C1	C1	B2	B2	B2	B2	B2	C1	B2
F/Spanish	C1	C1	B2	B2	B1	B1	B1	B1	B2	B2
G/French	B2	B2	C2	C2	B2	C1	B2	B2	C1	C2

Looking across the columns, it is clear that lecturers rated themselves somewhat differently for professional and non-professional purposes. One difference was a tendency to rate their vocabulary range higher for professional than for non-professional purposes. Very little difference, however, is seen in the ratings for grammatical accuracy and phonological control. Unlike these categories, the ratings for coherence-cohesion and spoken fluency show quite a bit of variability. Overall, what the differences in ratings suggest is that assessing professional language with generic scales could result in a misleading picture of language competence for professional purposes. Similarly, it could be expected that assessing ELF with a scale using the NS as the reference point against which language proficiency is measured may also result in a false picture of language proficiency for professional purposes. These aspects of language assessment will be examined in the next sections.

## 8.2 Sufficient CEFR levels for lecturing engineering subjects

By self-assessing their English on the five CEFR scales, lecturers would need to think carefully about each descriptor. This stimulus served as the basis of discussion in the semi-structured interviews. Using the scales and self-assessments, I prompted lecturers to discuss how they had decided where to place themselves on each scale, whether they felt their choice represented an adequate measure of their professional English for lecturing, and what CEFR levels they considered to be sufficient for this purpose.

Table 8-2 presents the CEFR levels that lecturers considered to be sufficient for lecturing technical subjects as well as the level set by language and communication experts in the pilot-certification assessment (see section 2.1). In the left-most column, the table shows the descriptor levels, ranging from A1 to C2. The remaining columns show the names of the descriptor scales, which contain lecturers' ratings. In addition, the first two scales, coherence-cohesion and spoken fluency, are categorized as pragmatic competences in the CEFR, and the remaining scales as linguistic competences. The yellow highlight marks the minimum CEFR level set by communication experts for each descriptor. (See section 5.3 for descriptor scales.)

Looking across the columns, there is a clear difference between the CEFR levels recommended by EMI lecturers as a minimum requirement for lecturing than those recommended by the language and communication experts in the pilot-certification assessment. On the whole, the minimum levels suggested by the lecturers are lower. For lecturing, they rated coherence-cohesion and vocabulary range at a higher level (B2) than phonological control (B1). On the remaining two scales, lecturers' recommendations for a minimum level varied. What is more interesting is why they chose these particular levels as being sufficient for lecturing. These findings are presented next.

**Table 8-2. Five CEFR scales: levels EMI lecturers consider sufficient for lecturing engineering subjects.**

	Cohesion & Coherence	Spoken Fluency	Grammatical Accuracy	Vocabulary Range	Phonological Control
C2					
C1		1			
B2	7	3	2	7	
B1		3	5		7
A2					
A1					

### 8.2.1 Telling a story logically is important and jumpiness is normal

The lecturers' descriptions of coherence/cohesion implied that 'jumpiness' in long contributions is normal and that telling a story logically is important. As sufficient measures of coherence/cohesion, both B2 descriptors were selected (see the coherence-cohesion scale in 5.3.3).

In discussing the cohesion-coherence scale, the phrase 'to tell a story' was used repeatedly to capture the essence of presenting lecture content in a logical progression. This phrase, however, was also present in the A2 descriptor, and thus may have been adopted. Terms pinpointed in the scales as descriptive features included 'a long contribution' and 'jumpiness', as shown in 8.1-.2.

- (8.1) <I> so you don't see any difference in your usage in professional and non-professional situations </I>  
 <D> well it depends *in free time* yeah because *you rarely have longer longer bodies of speech* that you produce it's usually a sentence two sentences at most </D><I> [mhm] </I>  
 <D> *it's a dialogue, it's not a story you tell* </D>  
 <I> yeah</I>

<D> so the logic the *logic comes in* comes in, well *you construct it together with the people you're discussing* </D>  
 <I> and how about when you're lecturing </I>  
 (...)  
 <D> yes *i rather often jump around when i'm speaking* i'm not sure it's bad actually but i do it quite much *i think i'm not very good at story telling in story telling you need to create a line not a line of argument but a line of a line of logical progression* </D>  
 <I> yeah i understand so b2 lower is good enough for lecturing </I>  
 <D> yeah yes i think so i think so *i think the logic of i'm often telling stories or i'm telling things that should be a story that aren't and then i get when i'm lucky i get a question from a student and precisely how does this work* </D>

The account implies that lecturing involves both monologic and dialogic discourse. The two types of discourse necessitate planned and unplanned speaking. The account also seems to imply that lecturing offers limited possibilities for co-constructed speech. Moreover, with a single speaker assuming the responsibility for the logical progression, it suggests that more 'jumpiness' is likely to occur. From the account, it appears that 'jumpiness' was viewed as serving a purpose: it may trigger a dialogic act – where a student asks a question that then restores the line of thinking or takes it in a new direction – a kind of co-constructed speech. The account suggests two ways in which lecturing differs from conversation, including turn-taking and logical construction. As the next account suggests, 'jumpiness' in a lecture is acceptable.

(8.2) <I> yeah </I> <E> [and] *jumpiness* well of course *jumpiness* is of course well *it's acceptable @@ if it's not every minute or so @@* </E>  
 <I> @@ yeah right @@ </I>  
 <E> @@ it says@@*even in a long contribution* so i really, *some jumps* in a lecture is not *doesn't sound very bad* </E>

The accounts (8.1-.2) describe lecturing as speaking for lengthy stretches, where jumping around occurs but should not happen excessively.

For coherence and cohesion, the skills described as belonging to professional language seem to resemble the descriptors at level B2 and lower. An examination of the scale suggests that descriptors move from features of informal conversational discourse in the lower end to features of formal discourse common to prepared written speeches or formal writing at the upper-end. According to the accounts, it appears that the lecturers' self-assessments of coherence/cohesion for professional English supports findings from a corpus-based study, which shows features of informal discourse as salient in ENL spoken academic discourse (Biber 2006: 70). The accounts also contrast with the prescriptive views of pragmatic competence discussed in 5.3.3.

All in all, the accounts suggest L2 lecture discourse to have features similar to ENL conversational discourse. Lecturers' perceptions of their professional English include



features such as ‘jumpiness’, seen as acts that may lead to co-constructed speech that help to repair or retrieve logical progression, as indicated in example 8.1.

### 8.2.2 Accessible (NNS) fluency is central

The fluency scale is explicitly NS-referenced, a point that made self-assessment ambiguous. For example, the account in 8.3 suggests this reference to be irrelevant to professional communication.

- (8.3) <I> all right on this spoken fluency chart where would you place yourself </I>  
(...)  
<I> all right er </I> <C> *[i th-]think this is a little bit crazy this reference to native speaker that's not, here it no-not the real question how well i can speak with na-native speakers ok @@ in so-some@@ occasions but there might be </C>*  
<I> so would this description be relevant for a lecturer here </I>  
<C> not exactly so, if i need some more (xx) *then it's technically to be able to communicate with someone non-native that's still, it's more relevant definitely </C>*

Unlike 8.3, the next account seems to accept the NS as a normal assessment criterion.

- (8.4) <I> and er let's begin with the same first two questions, where would you place yourself for everyday use and for professional use </I>  
<E> erm just something comes to my mind that i like, i'd like to say that *i think this er describing the interaction with a native speaker is <READING> possible without imposing a strain on either party </READING> i think that's er an excellent phrase to describe</E>* <I> [which one]</I>  
<E> that's b2 lower b2 it was already in the earlier place also and *i think that's a very nice that's a very nice er quality to or kind of conceptual description of this a strain on either party the emphasis on either party is quite nice so that also the mother tongue is the mother tongue speaker the native speaker isn't getting crazy with more talking with the persons @@ that's a good quality @@ </E>*

However, as indicated by ‘also the mother tongue ...’, the account seems to imply that “without imposing strain on either party” (Europe 2001:129) is a good quality for NNS spoken fluency, including NS-NNS and ELF interaction. The demands for achieving this aspect of communication, however, may change with the situation and role. As the next account suggests, fluency in everyday conversation does not generally require discussing conceptually difficult topics, whereas lectures do.

(8.5) <E> (...) but well it just came to my mind a very nice way of saying that, yes erm i would as earlier say that private life since i don't have *i don't talk so often about conceptually difficult subjects in private* <LIFE> what does that tell about me@@ but@@maybe rarely@@*but anyway i would say that it's* <SPOKEN FLUENCY> *smoother in private life, lecturer environment is always a bit difficult but i like the description that longer and er complex stretches of speech which in lectures for example is such a thing talking for 90 minutes* </E>

For lectures, the descriptors 'conceptually difficult' and 'lengthy utterance' are not surprising. As the account implies, changing the situation and roles places different demands on fluency. An additional problem is the assumption of a generic context, a point further discussed in the next paragraph.

Perceived fluency adopts a listener perspective, which consists of the impression the listener forms about the speech planning and production as functioning easily (Lennon 2000:27). In a cooperative exchange, such as a lecture, the goal of the speaker is to hold the attention of the listener in the given situation. In EMI lectures, the fluency variables include NNSs (primarily) and a teacher-student relationship. The NNS context means listeners need to deal with high variability in spoken language (due to a wide variety of L2 accents). The teacher-student relationship also means listeners must deal with cognitively demanding content. Given these contextual variables as well as the monologic nature of lectures, pausing is a necessary feature for comprehension, as the account in 8.6 suggests. In terms of fluency, what the contextual variables suggest is that the monologic character of lectures and the cognitively demanding content are more appropriate for determining the degree to which listener attention is held in lectures than the generic CEFR fluency scale anchored in NS-NNS interaction. For holding listener attention in lectures, the account in 8.6 identifies the key phrase "a fairly even tempo" (Europe 2001: 129) from the B2 (lower) descriptor as a desirable feature.

(8.6) <E> erm *so it does sometimes happen that i speak maybe not for 90 minutes but anyway a fairly long time before something something is discussed* or so, that's *i think that describes the the lecturer situation quite nicely* </E>

(...)

<I> would you also would that be the level you think would be appropriate as well for a lecturer </I>

<E> oh i think er this b2 lower <READING>*fairly even tempo* </READING> *and sometimes pauses*, ok i think i think that's *still acceptable in a lecture environment* so b2 lower </E>

<I> all right </I>

<E> *but then but then it would become too difficult for the students if it is b1* which sometimes i said even b1 is good enough but well it's different obviously </E>

Even though the B2 descriptor makes reference to regular interaction with NSs, the accounts (8.4-.6) do not suggest this skill to be necessary or relevant to lecturer fluency. Rather, the accounts pinpoint other descriptive phrases as important, such as some pauses and no strain imposed on either party, as well as discussing conceptually difficult topics. With the exception of maintaining 'a fairly even tempo', the key phrases seem to coincide with one of Fillmore's (1979) criteria for spoken fluency, coherence, which is defined as "the ability to talk in coherent, reasoned and 'semantically dense' sentences" (1979 [2000]:51, emphasis original). Here, 'semantically dense' could be perceived as conceptually more demanding, thus making it appealing for describing spoken discourse in lectures.

To gain insight into temporal variables, I asked lecturers to elaborate on their experiences with pauses in lectures. The accounts suggest that pausing occurs due to time required for managing lexical gaps and for conceptualizing ideas.

- (8.7) <E> <ON PAUSING> *but mostly it's really terminology i think it's technical terms lacking suddenly, that's it* </E>  
 <I> so how do you handle that </I>  
 <E> yeah well in different ways i suppose *sometimes the lecture slides help* </E>
- (8.8) <I> ok, so you were saying these gaps and slowing down </I>  
 <C> yeah. (...) i i don't know what is the process for me@@to find the words in English@@but sometimes even some very common words i don't just remember, ok what is yesterday i had this *i couldn't remember what is feedback er-ah-er-ah i was a little bit er@@disappointed in myself@@i i miss sometimes very common words in english* </C>  
 <I> yeah it's funny that those gaps sometimes come *so what did you do* </I>  
 <C> i i erm *i'm not sure i think i er gave them the paper* </C>

The accounts in 8.7-.8 suggest lexical gaps to include both infrequently and commonly used words. While the account in 8.7 implies common words to be suddenly missing, the response in 8.8 explicitly states common words as being forgotten. The accounts also suggest that lexical gaps include infrequently used words, such as terms for classroom management and specialty areas. Although lexical gaps occur, the accounts indicate that different strategies are employed for overcoming them. Unlike these accounts, example 8.9 suggests pausing occurs due to the need to conceptualize ideas in a way that makes them accessible to students.

- (8.9) <I> ok in what kinds of situations <DO YOU PAUSE> or for what reasons </I>  
 <A> er *sometimes just looking for the right word* </A>  
 <I> and how about other times </I>  
 <A> *sometimes thinking about the matter itself how to explain something which you just have an idea in mind but you still don't know how to explain it not in terms of language but in general, even if you are speaking your own language you would stop and think how should i explain this phenomenon*

now or this concept to the students then you have to think first of the logics of the thing and then put that logic in words then comes the language of course when you try to put it in words </A>

While lexical gaps may be L2-related, pausing linked to message conceptualization could apply to any speaker (Ln). Either way, some pausing for natural reasons seems acceptable in lectures. This view aligns with the literature on dysfluencies discussed in 5.3.3, which also shows the CEFR descriptors for pragmatic competence to be prescriptive and unnatural.

Central to lecturing is the ability to communicate ideas and to stimulate interest in the topics, as this next example implies.

(8.10) <I> er what do you feel that you need to be able to do to lecture well </I>  
<D> well of course *i have to be on top of the substance and be able to explain it in, not only in the correct way in a way that you know is professionally correct and right and so on but in a way that actually er that how how should i put it that that explains it that makes it clear by power of example by power of-of metaphors also the the real time aspect the fluency you need it to be able to create sentences in real time and then the lecture notes have to be good with lots of graphics and perhaps other materials too* </D>

Example 8.10 seems to suggest the main focus when lecturing to be on demonstrating adequate knowledge of the subject matter under the pressure of real-time processing, which entails articulating the content both in an accessible and engaging manner. In this way, the account implies that these aspects of spoken fluency are closely integrated into lecture discourse.

In all, the accounts seem to imply that spoken fluency descriptors for professional communication should target accessible fluency rather than drawing attention to native-like fluency. In order to achieve the goal of accessible fluency, it appears that the target level should include descriptors relevant to the ability to engage the listener. A suitable description would also consider situation, as fluency demands for everyday conversation differ from those for professional discourse. With lecturing, for example, the accounts seem to suggest that accessible fluency requires the ability to speak about content matter that may be conceptually difficult but which is presented in an accessible manner, where accessibility seems to refer to the ability to express ideas clearly through explanations and examples at an appropriate level of conceptualization. It also seems to include the ability to employ 'semantically dense' utterances without imposing strain on the listener and to pause naturally and accommodatingly.

### 8.2.3 A reasonable accent is fair

When asked to identify key words that guided their choice of a sufficient level of phonological control, the presence or absence of the term 'foreign accent' was central. The

highest descriptor in the scale that contained that term was B1, which lecturers deemed as sufficient for lecturing.

The absence of 'foreign accent' as a key descriptor in the upper half of the scale (B2-C2) seemed to imply that phonological control was approaching a native-like accent. This interpretation seems to have steered lecturers away from the higher levels. The view in 8.11 illustrates how the use of a term, such as foreign accent, introduces discrimination into a scale that should be neutral, instead.

- (8.11) <I> all right, i'm looking at this scale as well i'm just wondering what key words caught your attention </I>  
<D> erm yeah well <READING> *clearly intelligible* </READING> and <READING> *foreign accent* </READING> certainly yes they capture the essence </D>  
<I> ok and er what do these terms say to you </I>  
<D> *clearly intelligible means that people understand what you're saying but foreign accent well that's another matter it's a test of whether you've been able to overcome your origins and acquire a native control of the pronunciation* </D>  
<I> yeah </I>  
<D> yes of course yes but *here's also a sociological process at work here it's not just about speaking well or speaking less well* </D>  
<I> yeah </I>  
<D> *people will label you as a foreigner when they hear you speaking with a foreign accent and that may have all kinds of affects on the attitude towards you and that's really a social issue more than purely teaching well it's not directly related to that, it's it has these aspects it's it's you know erm being speaking about a foreign accent it's speaking about a one-of-us thing is he one of us or is he an outsider* </D>

As the example recognizes, the term 'foreign accent' carries sociological connotations related to being an outsider. Not only does such a term equate foreignness with intelligibility as a basis for assessing language competence, it also reflects a prejudice that should have no place in a testing scale. Yet, this term is characteristic of descriptors based on a NS model as the measure of success (or lack of it). Moreover, the next account questions the adequacy of this measure for ELF users, the majority of whom have been taught by NNSs of English and who have never lived in an English-speaking country.

- (8.12) <I> ok and do you think that measuring your own accent against a native speaker would be relevant for the use of english here for work </I>  
<D> not really no not really because *i think most of the students here have been taught english not by native speakers* </D>  
<I> (...) and how about *to be intelligible in lectures do you think that one needs to speak like a native speaker* </I>  
<D> no i don't think so *it might even get in the way* </D>

<I> ok </I>

<D> it might get in the way for instance in the way that *you start speaking too fast you get going you get enthusiastic and you forget to hold back and speak a bit slower and clearer* </D>

This view, for example, suggests that having a native-like accent may work against intelligibility in an ELF context. It implies that a rapid speaking rate negatively affects NNS lecture comprehension, a view supported by several studies (e.g. Flowerdew and Miller 1992; Griffiths 1990, 1992; Huang 2004). In contrast, other studies have found reduced speech rate to affect comprehension positively (e.g. Kelch 1985; Hynninen 2010), a view that supports speaking slowly and clearly. Rather than targeting native-like pronunciation, the account seems to suggest ‘clearly intelligible’ to be a good target for ELF users.

To be intelligible, speech perception should approximate that which an audience might expect, as this account suggests.

(8.13) <I> ok erm i was wondering about like er *to be successful at lecturing in english do you think it's necessary to have a native-like accent* </I>

<D> *no i think it might even be a problem it's i think you need a well accent accent you should let me say you should have a correct accent you shouldn't pronounce in a way that people have difficulty with it but native-like erm i don't think it helps as long as it's a reasonable accent a reasonable it sounds close to what the students expect* </D>

From the example in 8.13, it appears that a reasonable accent is determined by speech perception, as indicated by ‘you shouldn’t pronounce in a way that people have difficulty with it’. Given this qualifier, ‘a correct accent’ seems to suggest an ability to pronounce the sounds of English appropriately. According to recent research, appropriateness for ELF intelligibility would include core phonological items for English as a lingua franca (see Jenkins 2000). Moreover, as the account excludes a native-like accent as the best alternative, this view hints at NS accents containing features that are unintelligible to ELF users. This view also supports findings on ELF intelligibility (see Jenkins 2000). These views underscore acceptance of variation in accents, where intelligibility is the main concern. All in all, the view seems to support the notion of lingua franca intelligibility, an idea reminiscent of Jenkins’ lingua franca core (2000).

Even though the previous accounts suggest that native-like pronunciation can be disadvantageous in ELF contexts, the next account finds some features of native-like pronunciation to be desirable.

(8.14) <I> yeah ok er how about *your own target for pronunciation, what is it* </I>

<A> *it's not my my goal at all to try to speak like a native speaker maybe maybe some aspects the native speakers are using is some of my goals like like putting the stress in the right place in the word or change in the variation of stress in in a sentence to to make something more important than*

something *which i do in my native native language* and i like it very much because *you can express yourself twice better with with that than just with words which are coming monotonic* </A>

<I> yeah </I>

<A> so that's one thing i would like to learn and er *not to make let's say those pronunciation mistakes that makes the word not possible to understand by not only the native speaker but by anybody else (...)*</A>

According to the account, targeting correct word stress and utilizing sentence stress may aid comprehensibility. It is, however, unclear from this example what model of sentence stress is being employed: L1 French, L1 English, or a blend of L1 French and English. Regardless, the account suggests that targeting such features could increase comprehensibility, and this view receives some support from the literature. While the role of word stress is not yet fully understood in NNS communication, some studies indicate incorrect word stress to have a negative impact for NS and NNS listeners (Field 2005; Rajadurai 2006). Moreover, Jenkins (2000) identifies nuclear stress placement as critical to ELF.

On differences in pronunciation for professional and non-professional use, the following example implies that professionals make an effort to adjust their English in spoken interaction with NNSs.

(8.15) <I> okay so erm *but how is that <PRONUNCIATION> different for the everyday use and the professional use* </I>

<D> *you don't encounter as many different accents with the professional use people they they especially when they notice that i don't understand they switch to high english and generally they speak well americans speak american english and british speak british english and indians speak indian english it's all, but they sort of try to make it clear (...)* you know *i was working in denmark and the first three to four months i didn't get anything that was being said on television but then suddenly a switch flipped then because it's it's really a pronunciation thing* </D>

The account suggests accommodation to be a useful skill in international communication. It also implies that listeners adjust to accents over time.

In short, when self-assessing their English on this phonological scale that draws attention to distinctions between L1 and L2 use, the lecturers regarded NS targets as biased. It seems clear that the scale for phonological control targets SLLs whose aim is to acculturate to an ENL target culture. This aim entails acquiring phonological control that approaches native-likeness in order to blend in and be accepted. However, as the accounts reveal, it is not the goal of lingua franca users to acculturate to ENL norms – but perhaps to some other norm to which they have a close affinity. The accounts also present a view of accents needing to be ELF intelligible, and of speakers needing to be able to accommodate and adjust to different accents. The views suggest that the lecturers orient towards a neutral concept of phonological control, that is, being ‘clearly intelligible’ and

having a ‘reasonable accent’. These concepts seem to capture the essence of suitability in assessing ELF phonological control, and should not be confused with native-likeness. Moreover, these views of phonological control contrast sharply with the deficit views presented in the CEFR scale, criticized in 5.3.1.3.

#### 8.2.4 Non-native features are acceptable

On accuracy, one factor that may have influenced lecturers’ views is an article distributed in the mentoring program, which discussed speaker credibility. The article presents errors from a deficit view, maintaining that systematic errors (especially from the mother tongue) threaten NNS credibility and thus should be eliminated from their language<sup>24</sup>.

From the grammatical-accuracy scale, two levels were suggested as sufficient for lecturing technical subjects: B2 (upper) and B1 (upper). When asked about their perceptions of the overall scale, lecturers commented on the descriptors toward the top of the scale reflecting native-like proficiency. This impression appears to be related to the absence of the phrase ‘mother tongue influence’ in the B2 – C2 descriptors. Moreover, the account in 8.16 implies that mother tongue influence extends beyond the A1-B1 levels.

- (8.16) <I> and can you pinpoint something specific here that helped you determine your choice </I>  
<E> well the description of c1 and b2 upper is a bit similar, but i think there are there are this this er slips or <READING> minor flaws in the structure </READING> i would i mean i was thinking between or deciding between c1 and b2 upper and i chose b2 upper because it feels that or i feel that the slips and minor flaws are there and they are rare but not very rare </E>  
<I> all right </I>  
<E>and they are not difficult to spot *i think difficult to spot may be pending that the c level is more erm i feel that they are more for native native speakers maybe* </E>  
<I> and what made you makes you think that </I>  
<E>@@well@@er well obviously they are, maybe they are not from the concept point of view but the <READING> *consistent grammatical control of complex language* </READING> *is not easy to acquire as for a foreign language i think so that’s@@that’s just my analytical point of view@@* </E>  
<I> all right erm was there anything else that in the scale that er made you think that this upper level could be targeting a near native speaker</I>  
<E> *this b1 this mother tongue influence* </E>  
<I> yeah</I>

---

<sup>24</sup> I did not know about this article (i.e. Lavelle 2008) until it was pointed out in one of the interviews. Since it was distributed to all mentees, I assume they have all read it.



<E> *well that of course i'm aware that i have mother tongue influence but i so i didn't consider that strong enough to be@@@to put me to b1@@* </E>

- (8.17) <I> all right, ok, and now the title of this table grammatical accuracy erm what does this word accuracy mean to you </I>  
<D> grammatical accuracy, it means you er you don't make mistakes it's a very discrete thing you either make a mistake or you don't in this case for grammar </D>  
<I> ok erm so that's very clear erm what level on this scale do you think would be sufficient for lecturing in english to non-native speakers for example here in finland </I>  
<D> *in finland i think that b1 is actually the one in that case* </D>  
<I> and why would you choose that level </I>  
<D> erm *because a2<CEFR LEVEL> isn't good enough because then the errors you make are going to distract the listener* </D>  
<I> all right </I>  
<D> *with upper you i mean it's clear that you're not a native speaker but it doesn't harm it doesn't get in the way* </D>

The two accounts, 8.16-17, demonstrate the difference between a deficit view and a neutral view to grammatical accuracy. In 8.16, the account illustrates how deficit definitions can cause confusion for assessing the SLU of English: E recognizes that his grammar has mother tongue influence but considers his level higher than B1, thus chooses descriptor B2 (upper). In 8.17, the account does not seem to consider whether mother tongue influence extends beyond level B1, but does consider this feature to be acceptable for lectures and thus chooses this level. This difference raises the question of whether it is necessary to describe grammatical mistakes with deficit terms in the lower half of the scale, marking features from the mother tongue as undesirable. A critical discussion of this prescriptive aspect of the CEFR grammatical accuracy scale is presented in 5.3.1.1.

Another inconsistency that concerns the harmonizing of the scale is a shift in focus. According to the next account, the descriptors in the lower part of the scale focus on communication and those in the upper part on accuracy.

- (8.18) <I> (...) how is grammatical accuracy being judged here in the scale </I>  
<C> ok, yeah . (xx) how well it's understood, anyway on these levels <SIGHS> *on these levels <INDICATING B1 AND BELOW> i think it's more about the understandability but here on the higher levels more towards accuracy* </C>

Such a shift in the internal consistency in a scale can cause ambiguity in professional assessment, where communication is primary for accomplishing work-related tasks and accuracy secondary. An inconsistent scale can thus lead to an unreliable or invalid assessment, especially for SLU where the focus is on efficient communication not

accuracy. Likewise, the NS-referencing is also problematic as it is irrelevant to professional communication in an international context (8.19).

- (8.19) <I> ok, all right and er how about when you read through these would these grammatical descriptions would they be something that you could apply to your work situation </I>  
<C> <SIGHS> *ok i can evaluate myself based on these definitely this is in a way clear but this this is not necessarily exactly er relevant from my my work perspective* </C>  
<I> ok you said this is not relevant particularly </I>  
<C> not not exactly b-because *there is only one native speaker in my class anyways so it's er it's not so relevant what is the difference between my myself and native speakers they <THE STUDENTS> are not native anyways ok some of them may may be able to assess that aspect but (xx) what's the difference between native spoken english and er but in that situation when i'm giving the lectures that's not really the main point* </C>

On neutrality, the accounts in 8.16-.19 suggest that there is no reason to bias scales with NS norms. In other words, it is not the aim of professionals communicating in ELF to achieve native-like proficiency or to erase mother tongue influence from their spoken grammar (regardless of their level). These accounts provide further support to the irrelevance of native-likeness to the evaluation of professional communication.

On this point, one account differed, where the speaker explicitly refers to the NS-biased article mentioned in the beginning of this section. The example in 8.20 first identifies level B1 as adequate for lecturing, a descriptor that mentions noticeable mother tongue influence. However, the account then suggests this level to be insufficient for speaker credibility – according to the NS-biased article.

- (8.20) <I> and the last one grammatical accuracy what level do you think would be sufficient <FOR LECTURING ENGINEERING SUBJECTS> </I>  
<A> grammatical yeah, i would say that even b1 would be enough here but there is one aspect here the *credibility of of the teacher and if you want to have a good credibility* you probably need to have *at least b2 or c1* so that's i think it comes from that article <GOES TO GET IT; SEE EXPLANATION IN BEGINNING OF THIS SECTION> *so the more systematic mistakes you make in the classroom when you are teaching er the less credibility you get from the students* so so in some cases er the gram the grammar *i er find that the grammar is very important in language i think it's the basis of language the grammar if you if you don't handle the grammar well then you run into more problems actually* </A>

In this example, speaker credibility is viewed as a desirable trait for lecturing. It also appeals to the NS-biased article, and this speaker accepts that view. The account stresses the importance of good grammar and eliminating systematic mistakes as defining traits for

professional credibility. In contrast to this view, the following account contains no mention of the Lavelle article and links credibility to non-linguistic skills.

- (8.21) <I> ok erm how about credibility are you ever concerned about that when lecturing </I>  
<C> @@of course so i think it's important to have deep knowledge (xx) i should be able to answer questions but some sometimes these asian students are difficult to understand then then it's difficult </C>

From the accounts, the picture that emerges is that grammatical descriptors for assessing professional communication need to be consistent in their focus. The accounts also imply that some NNS features are acceptable, even at high levels of proficiency. Moreover, they indicate speaker credibility to be a desirable trait, and suggest the ability to demonstrate in-depth knowledge as important as well as having good grammar.

### 8.2.5 Specialized vocabulary is primary

The accounts on vocabulary for lecturing indicate idiomatic usage and colloquialisms to be problematic as measurement criteria. As a sufficient descriptor for lecturing, B2 was designated. It appears that this descriptor was chosen since it is the only one in the vocabulary-range scale with a key phrase that mentions vocabulary connected to field-specific knowledge. At the C1-C2 levels, the scale was not descriptive for professional purposes, including phrases such as having “a broad lexical repertoire ... [and] a good command of idiomatic expressions and colloquialisms” (ibid). Lecturer views to idiomatic expressions suggest them to be perceived as useful with NSs, not NNSs. An illustrative example includes the usage of idiomatic expressions at international conferences, where NNSs have difficulties comprehending them due to unfamiliarity.

- (8.22) <I> ok erm and would you say that it's important for you to be able to use idiomatic expressions and colloquialisms when you're communicating in english at work</I>  
<F> *well probably not at work because it's a more let's say er formal type of language based on on a technical technical issues so, as long as you are able to express those in a formal way with a sort of scientific type of expression then it should be ok but er for the daily usage of the language then yeah i mean to be more let's say to express yourself a bit better with er native speakers then probably this kind of colloquial expressions may be may be good, even because i'm i also consider that i experience sometimes in conferences that i try to do this sort of colloquial type of expression then i lose the audience so @@ they don't follow me @@ </F>  
<I> ok yeah </I> <F> [so in] </F> <I> [yeah] yeah you've noticed</I> <F> [so i] try to i mean er like er yeah formal language when expressing some scientifically concept or whatever otherwise then yeah technically i lose*

*them so @@ that's why i think that so ok er when you are expressing these kind of er technical stuff it's probably better to be accurate and precise use the terms and terminology they are use to so </F>*

When asked directly about their views on using idiomatic expressions in lectures, the responses suggest it best to avoid them with NNSs. The account in 8.22 also suggests that it is better to use formal language for explaining scientific and technical concepts than colloquial expressions. Moreover, at a general level, what the next example illustrates is a fallacy in the scale, which assumes that learners first acquire general language for everyday use before learning language connected to a professional domain.

- (8.23) <I> so erm on this next scale this table 2 <VOCABULARY RANGE> yeah again look at it, where would you place yourself here for english for everyday purposes and where for english for work </I>  
<F> <READING THE SCALE> for daily usage i would say B1 and for work probably B2 because sometimes i mean yeah yeah a lot of words are new words that i don't know, so i er need to look up in a sort of dictionary most of the time so i mean normally when when teaching i have a good understanding of all the terms and so on but there's always some sort of new new words when reading some books </F>

The account suggests that professionals may acquire a stronger foundation in terminology related to their fields than in words for everyday use.

An aspect of vocabulary not covered in the vocabulary-range descriptors includes the ability to utilize words from other languages to cover local items or concepts, when necessary. The following example touches on this point.

- (8.24) <I> is there anything you feel you have some uncertainties with </I>  
<D> in english </D>  
<I> yeah </I>  
<D> erm professional english </D>  
<I> yeah </I>  
<D> not really no professionally i think i speak professional english better than professional finnish but sometimes yeah *sometimes vo-vocabularywise i have er some erm describing some finnish things in english i have difficulty with that* </D>

Example 8.24 raises an issue that is not uncommon in ELF contexts. Plurilingual users are able to communicate ideas through code switching, mixing, or blending languages. For example, a lecturer may introduce a Finnish concept using the Finnish word when an English one is not available. This plurilingual skill is relevant to the shared repertoire of ELF users in a given community of practice, where it demonstrates a communicative ability. This position contrasts with the prescriptive CEFR descriptors for lexical competence, as discussed in 5.3.1.2.

For assessing the vocabulary of lecturers, another aspect missing from the vocabulary-range descriptors includes cross-disciplinary domains. The next examples, 8.25-.26, draw on the importance of having a broad range of vocabulary for cross-disciplinary purposes.

- (8.25) <I> do you feel that you need a broad repertoire of words in your subject area</I>  
<G> in in <NAME OF ENGINEERING FIELD> yes i think because it's it's quite wide </G>  
<I> ok that's a little different from some of the other areas of engineering </I>  
<G> erm it's *it's a kind of field which is er let's say border with many scientific and technology domains* </G>  
<I> right that's a good point </I>  
<G> *and we need er somehow to make a synthesis of erm of different scientific domains* </G>
- (8.26) <I> all right erm ok i guess one question about this one that erm i was wondering that in terms of vocabulary what would the word error means to you </I>  
<G> an error would be er <COUGHS> mixing of some concepts and so using the not not the appropriate terms to describe a concept the in engineering design a bit like design there is a course of thinking and this course are have somehow developed their own set of specific vocabulary and erm i've noticed if i use a word in from this school i cannot mix it with another word coming from the others <REFERRING TO SCHOOLS> because then the student will be totally lost and i have students coming coming from taik <MEANING AALTO SCHOOL OF ARTS> they they are using a set of terms that i don't manage very well erm i'm more using the set of terms used in engineering design and er so so sometimes when they are try to discuss with me taik <ARTS> type of students er they do not understand well what they are meaning sometimes @@ </G>

From these accounts, it appears that the measure of vocabulary range for lecturers should include knowledge related to cross-disciplinary fields in the upper descriptors as a measure of professional language.

All in all, what the accounts suggest is that the scale for vocabulary range centers on describing a range of vocabulary for general language ability rather than a range for professional language use. In addition, the range in this scale moves from general vocabulary (A2-B1) to professional (B2) to NS acculturation (C1-C2). This learning sequence builds on the fallacy that general vocabulary is learned first – as a core – onto which professional vocabulary is added. For SLU, this may not be the case at all. In short, it appears that this scale targets SLL that includes vocabulary for everyday use in an ENL context, not SLU that involves using vocabulary for professional purposes in specific or cross-disciplinary domains in an international context.

## 8.3 CEFR limitations

From the point of view of this investigation, the CEFR scales have limitations for both LSP and ELF. This section presents its shortcomings in relation to these two aspects.

### 8.3.1 Limitations for LSP

For assessing LSP, the CEFR scales have limitations, a point raised in previous literature (e.g. Huhta 2010; Toepfer and Virkkunen-Fullenwider 2005). A central problem is the lack of focus on professional vocabulary and discourse features for professional communication. Contributing to this discussion, Table 8-3 summarizes five limitations of the CEFR scales for assessing LSP.

The left-hand column presents five categories into which the limitations have been grouped. The right-hand column outlines the limitations. One shortcoming includes the treatment of assessment in generic terms, with descriptors based on

**Table 8-3. Some CEFR limitations for assessing LSP.**

CEFR Limitations for LSP	
Treats assessment in generic terms	➤ Limited to general language abilities of a 'learner'
Excludes situational appropriateness	➤ No consideration of appropriateness of language & communication skills for a given situation or role
Minimizes professional vocabulary	➤ Limited to B2 descriptor ➤ No measures for range of domain-specific vocabulary
Overlooks communities of practice	➤ No descriptors for discourse features common to communication among professionals (in a given community of practice)
Lacks indigenous assessment criteria	➤ No criteria for setting field-specific, professional standards

predetermined and fixed norms that are directed at a 'learner'. This approach lacks descriptors that conceptualize the social context in which an assessment occurs. The de-contextualized approach omits situational factors relevant to communicative competence, excluding facets such as situation, role, and appropriateness of language and communication skills relevant to determining the successful completion of professional tasks. In addition, the scale for vocabulary range lacks harmonization, assuming a wide range of categories that do not reflect measures for capturing the range of professional vocabulary acquired. Similarly, the scales do not consider communicative factors related to communities of practice (Huhta 2010). For example, what features of discourse are important to delivering a lecture successfully? On measurement for LSP, the literature suggests applying indigenous assessment criteria (e.g. Jacoby 1998), a feature missing from these scales. These limitations illustrate the inadequacies of the CEFR scales for assessing LSP.

For lecturing, the accounts suggest the descriptors are missing some aspects that would be relevant to assessing language competence for this target purpose, as summarized in Table 8-4.

In the table, the left-hand column shows four categories into which the missing aspects have been grouped. The right-hand column outlines the aspects drawn from the interview data that appear to be missing from the CEFR scales. The first category, speaker credibility, was raised as important for lecturing, where having good grammar is relevant, being knowledgeable about the subject matter is central, and having an ability to answer questions is pertinent. The second category, engaging the

**Table 8-4. Some aspects to consider for assessing lecturing.**

Some criteria to consider for assessing lecturing	
Speaker credibility	<ul style="list-style-type: none"> <li>➤ Appears (fully) knowledgeable on the topic</li> <li>➤ Handles questions well</li> <li>➤ Has good grammar</li> </ul>
Engaging the listener	<ul style="list-style-type: none"> <li>➤ Shows enthusiasm for and interest in the subject</li> <li>➤ Pauses appropriately, such that it enhances listening</li> <li>➤ Interacts with listeners</li> </ul>
Making content accessible	<ul style="list-style-type: none"> <li>➤ Communicates cognitively demanding ideas in a way that is easy to understand               <ul style="list-style-type: none"> <li>➤ Expresses ideas clearly</li> <li>➤ Explanations and examples targeted at appropriate level and enhance conceptualization</li> </ul> </li> </ul>
Linguistic aspect	<ul style="list-style-type: none"> <li>➤ Uses correctly specialized terminology</li> </ul>

listener, arose as important to the professional context, where monologic discourse is prominent. Aspects of engaging the listener include showing enthusiasm for the subject as well as pausing accommodatingly and naturally in a way that enhances listening. The next category, accessible content, is essential to lecture comprehension, especially when content is cognitively demanding. This category refers to the ability of the lecturer to present challenging content in an accessible manner, where accessibility means expressing ideas clearly through explanations and examples at an appropriate level. The third category, linguistic aspect, includes the ability to use correctly specialized terminology. These findings indicate the focus of the five CEFR scales to be too general for assessing LSP, as the scales lack performance criteria that include contextual aspects specific to the target purpose. Overall, the findings point toward the five scales as being inadequate for assessing professional English.

### 8.3.2 Limitations for ELF

For assessing ELF, the CEFR scales have some major limitations, a point also raised in earlier critiques (e.g. Ahvenainen 2005; Hynninen 2006; Seidlhofer 2003). Particularly problematic is the NS as the primary measure of language competence. This ‘learner’ goal is contrary to lingua franca usage, which primarily entails interaction among NNSs using English as a tool to accomplish tasks. Although the term ‘native speaker’ does not appear in every CEFR scale, the accounts suggest the scales were interpreted as NS based.

Related to the NS limitation is the notion of acculturation to a NS target culture, where monolingua-culturalism appears to be the objective. Although the CEFR acknowledges the need for competence in plurilingualism and pluriculturalism (Europe 2001: 4, 168), its scales concentrate on language competence for a target culture of native speakers, as emphasized by NS referencing. Such criteria for language assessment fail to consider the need for plurilingual and pluricultural skills in lingua-franca contexts. While this approach to language assessment may suit other European languages, it is ill fitting for English as a global lingua franca, which is comprised of more L2 than ENL speakers, who communicate in English outside a monolingual ENL world.

For spoken ELF, the following aspect seems to be missing from the descriptor scales: assessment criteria for a competent international SLU. The accounts suggest that suitable criteria for SLU would exclude a NS-bias. In other words, they would be neutral, reasonable, and fair for usage of English in international contexts. A related aspect, implied in the accounts, is the concept of speech being co-constructed and highly variable.

**Table 8-5. Some suitable criteria for ELF – missing from the CEFR.**

Missing from the CEFR: Some suitable criteria for ELF	
Fluency	<ul style="list-style-type: none"> <li>&gt; Accessible NNS fluency               <ul style="list-style-type: none"> <li>&gt; Where the NNS features enhance ELF communication</li> </ul> </li> <li>&gt;Accommodation and adjustment skills               <ul style="list-style-type: none"> <li>&gt; That enhance ELF communication</li> </ul> </li> </ul>
Accent	<ul style="list-style-type: none"> <li>&gt;Intelligible accent (whether NS or NNS)               <ul style="list-style-type: none"> <li>&gt; That represents an accent that could be expected by the intended ELF users</li> </ul> </li> </ul>
Grammar & Lexicon	<ul style="list-style-type: none"> <li>&gt; Some NNS influence               <ul style="list-style-type: none"> <li>&gt; That does not hinder ELF communication</li> </ul> </li> <li>&gt; Plurilingual strategies               <ul style="list-style-type: none"> <li>&gt;That enhance ELF communication</li> </ul> </li> </ul>

As shown in Table 8-5, one criterion appears to be NNS fluency that is accessible to ELF users. In other words, fluency is determined by the degree to which NNSs interact successfully to achieve a given task. Another criterion seems to be a reasonable accent, which was defined as one that is intelligible (regardless of the accent being native or non-native). A third criterion was the acceptance of some NNS features in lexicogrammatical constructions. The accounts suggest that NNS language is not perfect. They also suggest the use of plurilingual strategies to be a positive attribute in ELF communication.

### 8.4 Conclusion

In this chapter, I set out to investigate the extent to which five CEFR descriptor scales used in a pilot-certification assessment were suitable for evaluating professional English in an international context. The data were gathered during interviews with EMI lecturers, who were reading and interpreting the scales. The findings from the data analysis suggest the descriptor scales do not target skills relevant to the spoken professional English of



lecturers in a lingua franca context, which lends support to H2. Support for H2 was already evident in lecturers' self-assessments (in section 8.1), which pointed to different evaluations of their professional English and general language skills. Section 8.2 also revealed a clear difference between lecturers' and (language and communication) experts' perceptions of a sufficient CEFR level for EMI lecturing, as well as the reasons for lecturers' choices and what they considered to be adequate measures of their spoken professional English.

With regard to lecturers' perceptions of the suitability of the CEFR scales for measuring their professional English, the key findings are summarized here. In regard to the *coherence-cohesion scale*, the accounts suggest L2 lecture discourse to be perceived as similar to conversational discourse, a view that aligns with previous literature on L1 lecturing (e.g. Biber 2006). The accounts also view some 'jumpiness' in lecture discourse to be normal and telling a story logically to be important. The view of jumpiness included a positive function, perceived as an act that may elicit a dialogue that serves to restore logical progression. This view suggests some descriptive phrases in the lower half of the scale to be more suitable for describing lecturing than the upper half. In relation to the *fluency scale*, the accounts imply accessible (NNS) fluency to be central to professional communication rather than native-likeness. For lecturing, this fluency involves engaging the listener and presenting content accessibly. These criteria include an ability to present conceptually difficult content at an appropriate level using clear examples and explaining ideas clearly, as well as pausing naturally and accommodatingly in the given context. Here, 'naturally' may mean pausing due to an occasional lexical gap that is repaired in a non-distracting manner, for example. In other words, 'pausing naturally' is not described as native-like, but rather in terms of accessible NNS fluency.

On *phonological control*, the scale itself draws attention to distinctions between L1 and L2 use. The accounts viewed NS targets as biased, and a reasonable, intelligible accent as fair. In short, intelligibility was stated as the goal, not native-like English. As the accounts also rightly point out, scales that pit 'intelligibility' against 'foreignness' as a measure of communicative success (as this scale does) display overt prejudice. This concept has no place in language testing scales. Similar to the previous scale, accounts related to grammatical accuracy also viewed the scale as NS-biased. The accounts also imply some NNS features to be acceptable at high levels of proficiency. Good grammar was viewed as important for speaker credibility and so was in-depth knowledge of the subject.

The view of the scale for *vocabulary range* was that it is suitable for describing general language competence rather than professional language ability. The accounts also viewed the scale as applicable to SLL for an ENL context, but not to SLU where terminology for professional English is relevant in an international context.

The analysis also indicated some *limitations* in the CEFR scales *for assessing ELF*, a point previously critiqued in the literature (see discussion in 8.3.2). The central problem lies in the NS as the measure of language competence. The accounts suggest the scales to be interpreted as NS based, irrespective of whether they contain the term 'native speaker'. This term also represents a 'learner' goal of acculturating to an ENL target culture, a goal not shared by ELF users. This aspect of the CEFR renders it inappropriate as a tool for

measuring lingua franca competence. For assessing ELF, the present study indicates some suitable criteria as including accessible NNS fluency, an intelligible accent, accommodation skills, and plurilingual strategies.

*For assessing LSP*, the analysis also discerned several *shortcomings* in the CEFR scales, a point also raised in previous literature (see discussion in section 8.3.1). In the present study, the main drawback points to a lack of contextualization. More specifically, this drawback is characterized by the following: assessment of the general language abilities of a ‘learner’, exclusion of situational appropriateness, omission of communities of practice, minimization of professional (i.e. specialized) vocabulary, and absence of indigenous assessment criteria. For the purpose of assessing lecturers, the analysis also identified some aspects missing from the scales, such as speaker credibility, engaging the listener, making content accessible, and linguistic use of specialized terminology.

The present study also found the CEFR scales to lack harmonization relevant to assessing ELF, as follows: the scale for grammatical control is inconsistent in its focus on communication, and all five scales are NS-biased (whether implicitly or explicitly stated). The latter aspect was shown to cause ambiguity in self-assessing spoken fluency, lexicogrammatical ability, and intelligible pronunciation. For LSP, the scale for vocabulary range was shown to lack harmonization for assessing professional vocabulary.

## 9 Analysis of students' perceptions of English in lectures

As part of the triangulation strategies for this study, this chapter investigates students' perceptions of English in EMI lectures. It also answers the third research question, which examines students' perceptions of their ability to comprehend lectures in English and whether their perceptions change over the duration of a course. In order to address this question, data from a repeated survey were analyzed. For the investigation, the following hypothesis was formed: Students' perceptions of their lecture comprehension will change over time as they adjust to their lecturer's English. This directional hypothesis assumes a positive relationship. The results are presented in section 9.1, followed by discussion in 9.2, and then a summary in 9.3.

### 9.1 Results

For the repeated survey, background information about the participants is presented in 6.2.3. Table 6-3 also provides an overview of the language backgrounds of the lecturers (A-D) and their students. Briefly, the sample represents lectures where English is a true lingua franca with sixteen different L1s represented. Because the lecturers are viewed as competent SLUs of professional English, the results from the four lectures were combined to establish a general view. This section presents the combined results of the repeated survey, which has three sections (see Appendix F). The first section includes thirty-two Likert items (Q1-Q32), followed by two multiple-choice items (Q33, Q34), and then an open-ended item (Q35).

#### 9.1.1 Likert items

In this section, students' responses to thirty-two Likert items (Q1-Q32) from the repeated survey are analyzed. With respect to the presentation of the findings, the following should be noted:

- The total sample size reported reflects the number of paired students for the initial and post questionnaire ( $N = 36$ ) for all four lectures, where the sample size for each lecturer was as follows: Lecturer A ( $N = 8$ ), Lecturer B ( $N = 6$ ), Lecturer C ( $N = 17$ ), and Lecturer D ( $N = 5$ ).
- For the questions on a 4-point Likert scale, the mean values have been included. These mean values were calculated for each major theme based on the questionnaire scores. Given the nature of the rating scales (involving ordinal level), these mean values have one meaning only: they are one way of indicating to what

extent the students' perceptions of lecture comprehension differs between two points in time, that is, at the end of the first and last lecture in the course.

- The direction of the scales was corrected, where 1 = disagree and 4 = agree.
- Items that were reversed in meaning from the overall direction were also corrected. For these items, if the respondent gave a 1, it was changed to 4; if they gave a 2, it was changed to 3, and so on. For a list of the reversal items, see section 6.3.5 or Appendix F.
- For analyzing the three themes that form the basis of the questionnaire, the questions were grouped as follows:
  - Lecture comprehension: Q2, 5-7, 12-13, 18, 24, 26, 28, 32
  - Lecturer language: Q1, 3-4, 11-14-17, 19-21, 23, 25, 27, 31
  - Lecturer communication: Q8-10, 22, 29-30

For the analysis, I combined the data from the four lectures (N=36) and compared the scores on the Confidence in Students' Perceived Lecture Comprehension at Time 1 (the end of the first lecture) and Time 2 (the end of the last lecture). Therefore, mean score values were applied to two-factorial repeated measures analyses of variance (ANOVA) with time (first, last) serving as within-subjects factor and theme (lecture comprehension, language use, and communicative ability) serving as between-subjects factor. All analyses were performed using SPSS 20 (SPSS Inc., Chicago, IL). Testing revealed no differences for either the factor time,  $F(1,105) = .742, p = .391$ , or theme,  $F(2,105) = 1.41, p = .249$  (see Table 9-1 and 9-2). Due to the non-significant main effects, the observed reliable interaction will not be further discussed here.

**Table 9-1. ANOVA results of within-subjects contrasts.**

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
time	Sphericity Assumed	.054	1	.054	.742	.391
	Greenhouse-Geisser	.054	1.000	.054	.742	.391
	Huynh-Feldt	.054	1.000	.054	.742	.391
	Lower-bound	.054	1.000	.054	.742	.391
time * theme	Sphericity Assumed	1.196	2	.598	8.248	.000
	Greenhouse-Geisser	1.196	2.000	.598	8.248	.000
	Huynh-Feldt	1.196	2.000	.598	8.248	.000
	Lower-bound	1.196	2.000	.598	8.248	.000
Error(time)	Sphericity Assumed	7.613	105	.073		
	Greenhouse-Geisser	7.613	105.000	.073		
	Huynh-Feldt	7.613	105.000	.073		
	Lower-bound	7.613	105.000	.073		

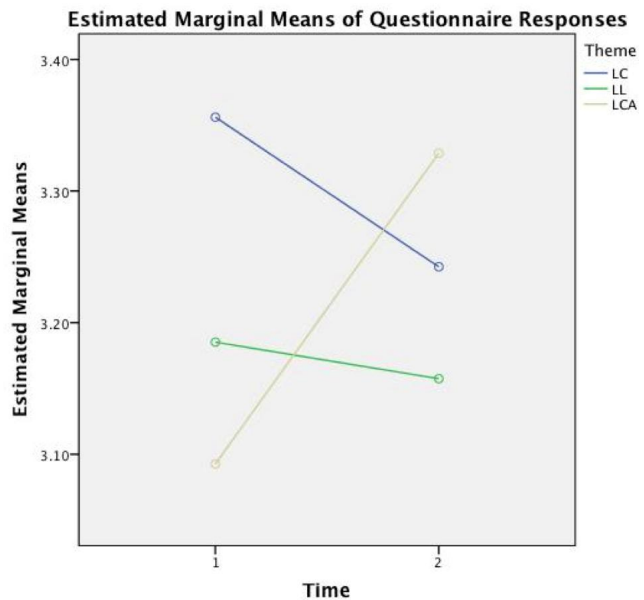
**Table 9-2. ANOVA two-factorial repeated measures results for tests of between-subjects effects.**

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	2249.409	1	2249.409	10255.616	.000
theme	.618	2	.309	1.410	.249
Error	23.030	105	.219		

Because the ANOVA results did not reveal very much about the relationships under investigation, the data were also examined for trends. To explore the marginal differences between the two time points, a scatter plot was generated. The differences are shown in Figure 9-1. In the figure, the x-axis shows time whereas the y-axis shows the mean score for the response values as measured on a 4-point Likert scale for each of three themes: lecture comprehension (LC), lecturer language (LL), and lecturer communicative ability (LCA).

Looking across all three themes, it can be seen that there are marginal differences in students' perceptions of lecture comprehension between Time 1 and Time 2. Figure 9-1 shows that students' perceptions of lecture comprehension and of lecturer English slightly decreased from Time 1 to Time 2 while perceptions of lecturer communication increased. The slight decline in the means for students' perceived lecture comprehension between Time 1 and Time 2 illustrates change over time although not adjustment to lecturer English. Nevertheless, the mean scores suggest that the overall comprehension was high across both points in time (with the lowest mean score being 3.25 out of 4.0). The high averages suggest a ceiling effect, which indicates that further investigations on this topic should use larger scales to better differentiate the positive range of responses.

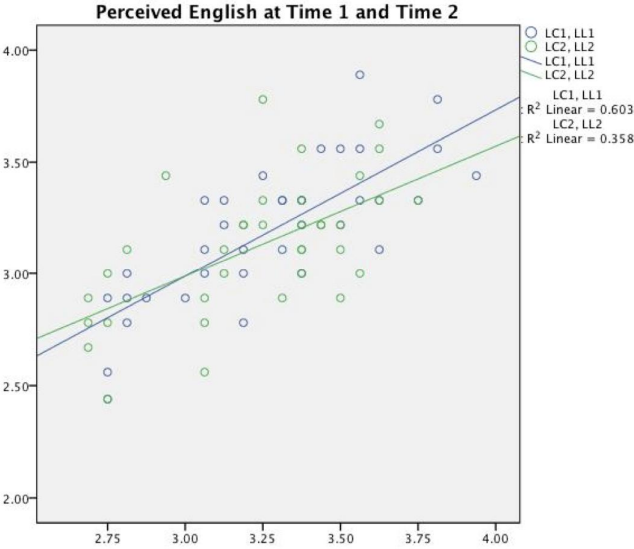
**Figure 9-1. Questionnaire mean response at Time 1 and Time 2 for each theme: Lecture comprehension (LC), lecturer language (LL), and lecturer communication (LCA).**



To further explore differences in students' perceptions of lecture comprehension, correlations were used to study the relationship between students' perceived lecture comprehension and their perceptions of lecturer English at the initial and final lectures (Fig. 9-2). To examine this relationship, scatter plots were generated and the squared correlation coefficient ( $r^2$ ) calculated.

In Figure 9-2, the scatter plot illustrates the linear relationship between perceived lecturer English and perceived lecture comprehension at Time 1 and Time 2. The x-axis represents the mean score for the response values to perceived lecturer English, while the y-axis represents perceived lecture comprehension. All responses were measured on a 4-point Likert scale. In the figure, the blue line shows the relationship between these two variables at Time 1, and the green line at Time 2.

**Figure 9-2.** A scatter plot showing mean values at Time 1 and Time 2 as measured on a 4-point Likert scale, with students' perceptions of lecture comprehension on the y-axis and their perceptions of lecturer English on the x-axis.



What can be observed from the scatter plot for Time 1 and Time 2 is that students' perceptions of lecture comprehension and of their lecturer's English positively correlate. The positive correlation means that students' perceptions of lecture comprehension changed between Time 1 and Time 2 in relation to their perceptions of lecturer English.

**Table 9-3.** Correlation between perceived lecture comprehension and perceived lecturer English at Time 1 and Time 2.

To estimate how much variance the two variables (perceived lecture comprehension and perceived lecturer English) share, the coefficient of determination was calculated. As shown in Table 9-3, a correlation was found at both Times 1 and 2.

Subject	Time 1	Time 2
Lecturer (N = 36)	$r^2 = .60$	$r^2 = .36$

At Time 1, students' perceptions of lecturer English helps to explain sixty percent of the variance in responses to perceived lecture comprehension, and at Time 2 slightly more than a third.

To determine the strength of the relationship, the interpretation guidelines follow (Cohen 1988:79-81), who suggests the following:

- Small  $r = .10$  to  $.29$
- Medium  $r = .30$  to  $.49$
- Large  $r = .50$  to  $1.0$

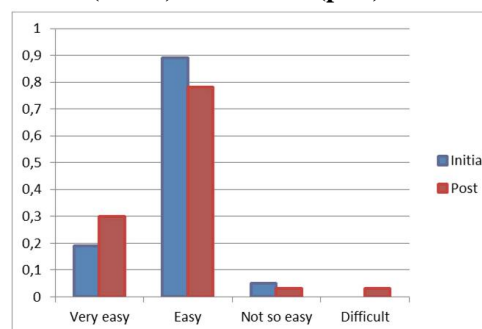
According to these guidelines, the relationships are strong at Time 1 and moderate at Time 2<sup>25</sup>. What the positive correlations indicate is that students perceived their lecture comprehension to be better when they perceived their lecturer’s English to be better. This finding implies that the lecturer’s English is important for comprehension at Time 1. This relation, however, is less important at Time 2. This difference means students adapted to the English of the lecturer or their comprehension becomes more disentangled from the English. This finding, which supports H3, should be further explored in future studies by having more sample times in the study than two.

### 9.1.2 Multiple-choice items

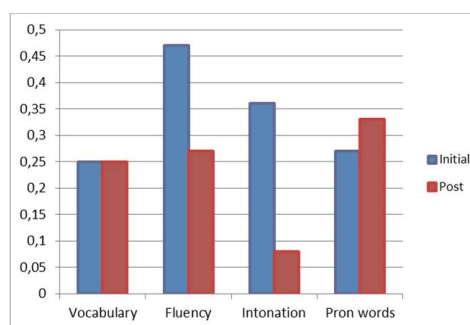
In Section II of the repeated survey, two questions (Q33, Q34) on lecturer English contained scales that differed from those in Section I. In Figure 9-3, respondents selected one item on the ranked scale in Q33 that corresponded to their perceived ability to follow the lecturer’s English at the initial and final lecture. What Q33 illustrates is the perceived ease in students’ ability to follow the lecturer’s English. While the results suggest that students found it easy to follow the lecture in both the initial and post survey, this ease was not without some challenges as shown in Figure 9-4. This chart presents aspects of the lecturer’s English perceived as challenging at two points in time, initial and post survey.

On Q34, some respondents selected multiple items as perceived challenges in the lecturer’s English on this categorical, non-ranked scale. At a glance, Figure 9-4 shows virtually no change on two aspects: vocabulary and pronouncing single words. However, a clear difference can be seen in perceived fluency and intonation, with students’ perceived challenges being much lower in the post survey than in the initial one. This difference suggests that students adjusted to these aspects of their lecturer’s English over the duration of the course.

**Figure 9-3. Q33-Students’ perceived ability to follow the lecturer’s English at Time 1 (initial) and Time 2 (post).**



**Figure 9-4. Q34-What students perceived as challenging in the lecturer’s English at Time 1 (initial) and Time 2 (post).**



<sup>25</sup> On determining the strength of the correlation, different authors suggest different interpretations of the value between 0 and 1. For example, in contrast to Cohen (1988), Jaeger (1990) suggests the following guidelines: below .30 = small, 30-70 = moderate, 71-90 = large, and above .90 = very large.

### 9.1.3 Open-ended item

In Section III of the questionnaire, students responded to one open-ended question. These responses provide qualitative data that supply further insight into students' perceptions of English in lectures. The questionnaires did not contain many responses, and thus are presented below as originally written. They are also grouped by theme and sorted within theme by initial and post survey responses.

Students' comments on perceived weaknesses in the lectures primarily point to delivery skills. Their comments relate to both lecture slide usage and pace, as shown in Table 9-4. These comments could pertain to any lecturer, not just EMI lecturers. The comments suggest improvements for further enhancing lecture comprehension, which is already quite high according to the quantitative analysis in section 9.1.1.

**Table 9-4. Students' comments related to lecturer delivery skills.**

Students' comments on lecturer delivery skills	
B/S2(L1-French)	The teacher may try to give more examples and not to stick to the slides. (initial survey)
A/S1(L1-English)	Sometimes, the lectures went on for too long, until the focus is off. It is rather hard to grasp important concept when the slides are just pictures with very few words to describe. Perhaps to improve, could add more description to a couple with his explanation. That would make the lecture much easier to follow. (post survey)
C/S3(L1-Chinese)	More clear slides could help to get the idea of the lecture when having problems to get understand the lecturer's English. (post survey)
C/S5(L1-Chinese)	The lecture of the teacher is fine and detail. I have no problem with the lecture. However, the speed of the lecture could be a little bit fast. (post survey)

On delivery of lecture content, students' comments also indicated perceived weaknesses. In Table 9-5, their comments offer a common-sense view, signaling content familiarity and complexity as factors affecting comprehensibility. While one comment indicates the advantages of content familiarity in lecture comprehension, the others indicate the disadvantages of unfamiliar content, including terminology. More specifically, the negative comments indicate that content atypical to their fields of study is more challenging to comprehend as well as unfamiliar terminology. In addition, content with complex aspects that may not be straightforward (i.e. 'tricky') also create challenges with lecture comprehension. Like the previous comments, these perceived weaknesses could be present in any lecture, whether EMI or not.



**Table 9-5. Students' comments related to lecture content.**

Students' comments on lecture content	
B/S1(L1-German)	To me, the lecture was easy to understand, maybe, I had a look at the course homepages before at home. So I had an idea of the course content before. But anyway, the lecture was easy to follow. (initial survey)
C/S1(L1-Czech)	This is not a typically [sic] technical course. (initial survey)
C/S2(L1-Amharic)	The difficulty in understanding the whole content of the subject may have arised [sic] due to the subject itself. I have taken previous course from this lecturer, and it was very easy to follow. However, human behavior modeling [the topic of this course] is more of physiological study than engineering. (post survey)
C/S3(L1-Chinese)	More time spent on explaining terminology would be also valuable for students to get know better of the lecture. (post survey)
C/14(L1-Chinese)	I'm interested in topics, but I hope to have more contents and theory, story ... just to make it more interesting because the name of the course is big and interesting. (post survey)
C/S19(L1-English)	Good course and I believe the lecturer has potential to develop himself more as he gives the course more in future years. (post survey)
D/S1(L1-Finnish)	Subject was rather tricky which influenced this. (post survey)

Comments related to pronunciation were ambiguous, as shown in Table 9-6. The comments simultaneously reflect student satisfaction and dissatisfaction with their lecturer's spoken English.

**Table 9-6. Students' comments related to pronunciation.**

Students' comments related to pronunciation	
A/S2(L1-German)	I am happy with the general manner this lecture is held. But I remarked that the lecturer did not make big progression in pronouncing some English words (for example, frame). In my opinion, that's the main point where improvements could be done. (post survey)
A/S4(L1-Urdu)	The lecturer's English is quite good. He is the second teacher (whom I have met) whose English is very good and understandable. The thing which I feel is that intonation can be a little bit distracting for students. But my feedback is good about most of the courses regarding understanding the English language.(post survey)
C/S4(L1-Chinese)	I am satisfied with the lecturer. The Finnish lecturer's pronunciation is not so good, but doesn't effect us to understand the lecture. (post survey)

For instance, each comment contains a phrase indicating students are satisfied with their lecturers, yet they all mention an aspect of pronunciation that could be improved. Of these, one comment suggests that the lecturer's (French influenced) intonation was slightly distracting.

Comments on perceived strengths in EMI lecture comprehension include ease of comprehensibility, as shown in Table 9-7.

**Table 9-7. Students comments related to lecture comprehensibility.**

Students' comments related to comprehensibility	
A/S4(L1-German)	Generally, I am very happy wit this lecturer. Compared to other non-native English speakers, he explains very clearly. (post survey)
A/S6(L1-Korean)	I think I need to study English more if I cannot understand lecturer's saying. The lecturer gives us a nice lecture which I can understand well. (initial survey)
C/S6(L1-Finn)	The lecturer's English is easy to follow. It may be that because my native language is same as the lecturer's it makes it even easier. (post survey)
C/S7(L1-Swed/Finn)	Easy to follow at least for someone who understands Finnish. (post survey)
C/S8(L1-Finn)	I am so used to 'business English' that it doesn't bother or hinder me when listening to the lecture. (post survey)

The comments relate to three aspects: previous ELF experience, own level of English, and L2 listeners and speakers sharing the same L1. Moreover, comments related to interlocutors sharing the same L1 reflect findings from previous research. This point is further discussed in section 9.2.

In addition, other comments were more general and reflected positive views to EMI lecturing, as shown in Table 9-8.

**Table 9-8. Students' comments that reflekt English as a lingua franca views.**

Students' comments reflecting ELF views	
A/S3(L1-Korean)	Because he is not native speaker, his English can not be perfect, however he speak English well, understanding me and other students perhaps. I think his English is well enough to carry out these kind of lectures. (post survey)
B/S2(L1-Estonian)	English is nowadays like a third native language – needed everytime [sic] and everywhere. (post survey)
C/S9(L1-Nepalese)	I believe change is a gradual process and would like to state that teaching in English is a step towards the right direction. (post survey)

Although the comments mention that EMI lecture language may not be perfect, they also suggest that the English used is good enough and that EMI lectures are a positive step forward. Moreover, these views seem to capture the essence of students' overall perceptions of their EMI lectures, and they do not appear to be guided by standard language or NS language ideologies.

Generally speaking, the perceived weaknesses mentioned in students' comments could be present in any lecture. Even the comments related to pronunciation reflecting students needing to deal with variability in accents could pertain to a lecture where an L1 lecturer has a different dialect from his/her L1 students. In spoken language, variability in accents will always be present. The next section discusses this point from the perspective of foreign accentedness.

## 9.2 Discussion

While many studies have been conducted on L2 lecture comprehension in English from various points of view (e.g. Camiciottoli 2005; Flowerdew 1995; Huang 2004; Huckin and Olsen 1990; Jones 1999; Lee 2009; Miller 2009; Morell 2004; Mulligan and Kirkpatrick 2000; Strodt-Lopez 1991), very few have been performed on students' perceptions of EMI lecture comprehension (Airey 2009; Hellekjaer 2010; Suviniitty 2010; Suviniitty 2012) and the concept of EMI students adapting to their lecturer's English over time has been largely uninvestigated. For the present study, a hypothesis was generated from the assumption that the human capacity to comprehend spoken language is highly flexible and that this capacity also applies to advanced SLUs of English who regularly use ELF. The hypothesis that students' perceptions of their EMI lecture comprehension will change over time as they adjust to their lecturer's spoken English was partly supported.

Using repeated measures ANOVA resulted in no reliable statistical significance for lecture comprehension in relation to time or theme. For this reason, the data were further examined using scatter plots. An examination of the scatter plots for trends in the marginal differences of means for time revealed slight changes in students' perceptions of their lecture comprehension. This slight decline in the means lends partial support to H3 in that it shows change over time but does not indicate adjustment to English in the lectures. The high mean scores, however, at both points in time suggest that students comprehended lectures well.

Examining the association between students' perceptions of their lecture comprehension and of their lecturer's English through correlations revealed a positive relationship that provides support for H3. The correlation suggests that the lecturer's English became less important for lecture comprehension over time. This finding is interesting and deserves more exploration. Future studies should include more sample times in order to establish the point at which perceived adjustment occurs.

Both the high mean scores and the positively correlated findings lend support to research on EMI lecture comprehension indicating few, if any, problems (e.g. Airey 2009; Hellekjaer 2010). Because the correlated findings in the present study do not make a distinction between language form and function, it was not possible to compare them to those of Suviniitty (2012) or Björkman (2010). Unlike these studies, the present study assessed lecture comprehension of the lecturer's English as a unit rather than directing attention to language form and function separately.

Although L2 English is a common, real-world phenomenon that presents a source of speech variability, the lack of significant findings from the ANOVA analysis suggests that proficient EMI students either adjust rapidly to their lecturer's spoken English or they require very little adjustment. In either case, the responses to Q34 in Figure 9-4, nevertheless, suggest that these highly proficient SLUs did not adjust to all aspects of their lecturer's English (also highly proficient). For example, the mean scores for intonation and fluency reflect adjustment, while the mean scores for vocabulary and pronunciation of single words do not. What might be the reasons for the apparent adjustment to some aspects of the lecturer's English but not to others?

While common sense suggests that more variability would be present in spoken ELF than in ENL, a study by Nash (1969) based on casual observation claims that NNSs are more intelligible to other NNSs than to NSs. This observation could support the moderate to high lecture comprehension mean scores in the present study. More recently, van Wijngaarden (2001) and van Wijngaarden et al. (2002) also provided evidence that supports proficient L2 listeners finding sentences produced by proficient L2 speakers to be just as intelligible as those produced by L1 speakers. The findings from their study also indicated an intelligibility advantage between an advanced L2 speaker and a novice L2 listener who was at a relatively early stage of L2 acquisition. These findings suggest the target language proficiency of the listener and speaker to be important for determining comparable intelligibility.

Because fluency, intonation, vocabulary, and pronunciation are intertwined with accent, the findings from the present study are compared to studies related to adjusting to accents perceived as different. On adjusting to accents, Chambers and Trudgill (1998: 3-4) argue that mutual intelligibility is related to listener exposure, willingness to understand, and ability to overcome deviations from language that is familiar or standard. Conversely, accents that are unfamiliar may cause listener irritation as well as listener judgments – whether the accent is native or non-native (*ibid*). These views involving NS-NNS or NS-NS interaction, however, do not align with findings from ELF studies involving academic language. For example, a study on EMI engineering students at a technical university in Sweden found ELF communication to occur with very little overt disturbance (Björkman 2010). In the present study, students' comments reflected tolerance towards NNS accents and attributed ease of lecture comprehension to some aspects of the lecturer's NNS accent, such as a shared L1 with the lecturer and exposure to BELF. On L2 speakers sharing the same L1, previous studies show that accent familiarity has a facilitating effect (e.g. Adank et al. 2009; Bent and Bradlow 2003; Sajavaara 1986). BELF studies also indicate tolerance for variety in spoken and written communication, and describe ELF as an accepted, normal practice in business communication (Ehrenreich 2009; Kankaanranta and Louhiala-Salminen 2010).

When interacting with foreign-accented speech, listeners may need to handle noticeable deviations considered to be non-standard in a target language. To process this type of speech, common sense suggests that listeners with extensive contact with NNSs would be more skillful at comprehending foreign-accented speech than listeners with little or no exposure. To date, it appears that this wisdom has not been studied in ELF interaction. For this reason, this discussion includes findings from NS-NNS interaction. A recent study on rapid and long-term experience with foreign-accented speech shows that mispronunciations of lexical items did not create lexical ambiguity for Dutch speakers listening to German-accented words (in Dutch) and that long-term experience facilitated word recognition with large and small acoustic deviations while limited experience (i.e. four minutes of prior exposure) facilitated recognition of small acoustic deviations (Witteman et al. 2010). At the level of word recognition, another study on interlocutor effects on the degree of perceived foreign accent found speech perception to change in a very short period of time: within one minute (Clark and Garrett 2004), and in less than five minutes (Flege and Fletcher 1992). At the word-recognition level, Bradlow and Bent

(2008) also showed faster adaptation to relatively highly intelligible foreign-accented speech than to relatively low intelligible foreign-accented speech. While these studies involved NS-NNS interaction, they provide support for highly flexible speech perception, showing rapid adaptation to spoken language that deviates from L1 speech patterns. To what extent this adaptability carries over into highly proficient NNS-NNS interaction appears to be unknown.

Recent studies on lecture comprehension suggest challenges with terminology and lecture content to be common to both L1 and L2 speakers. For example, a recent study compared L1 and L2 lecture comprehension at three Norwegian (sample size 364) and two German (sample size 47) institutions of higher education using self-assessment scores (Hellekjaer 2010). This study found difficulties in lecture comprehension with key terms and concepts due to unfamiliar vocabulary and difficulties with distinguishing meaning of some words, regardless of whether the instructional language was L1 or L2 English (*ibid*). The study concludes that difficulties already present in L1 lecture comprehension may further aggravate L2 lecture comprehension. It also emphasizes that L1 lecture comprehension is not perfect. Similarly, a study on undergraduate physics students in Sweden found no difference in the effects of instructional language, L1 Swedish and L2 English, on the lecture comprehension of students who were not first-year students (Airey 2009). In other words, only first-year students found it to be more challenging to study in L2 English than in L1 Swedish, reporting difficulties with note-taking and passivity in asking questions. Hellekjaer's (2010) study also demonstrates proficient ELF users comprehending lectures well and underscores the dangers of assuming perfect lecture comprehension in the L1. Like his study, the present study also found some challenges with vocabulary in lecture comprehension (mean score 0.25 in initial and post survey), and responses to the open-ended question suggest concern over unfamiliar content and terminology. In comparison to these other studies, a weakness in the present study is the lack of comparable lecture comprehensibility in the students' L1s. On the other hand, comparable data for sixteen different L1s would have been difficult to obtain.

From the present study, students' comments also suggest vocabulary and content atypical to students' fields of study to be problematic. For instance, one lecture series in engineering from the current study handled topics from the social sciences, thus introducing content from fields unrelated to engineering. On this point, students' views to difficulties in lecture comprehension align with those from previous SLA studies on L2 lecture comprehension of subjects in different fields. For example, Miller (2009) found that the language used in different types of lectures had an impact on the overall lecture comprehension, where engineering students reported fewer problems with lectures on science and engineering than on humanities-type subjects. Flowerdew (1995) also points out that lack of relevant background information causes difficulties in lecture comprehension. This information includes examples or situations that pertain to a culture (e.g. Finnish) with which students may be unfamiliar.

While vocabulary and content appear to present some challenges in lecture comprehension, the moderate to high mean scores for perceived lecture comprehension at Time 1 and Time 2 in the present study suggests that students nevertheless managed quite well. This finding also aligns with the findings from previous studies related to lecture

comprehension in EMI programs, which have demonstrated that advanced SLUs are proficient in lecture comprehension and that they encounter challenges similar to those of L1 speakers.

### 9.3 Summary

This chapter investigated whether students' perceptions of their lecture comprehension changed over time, using ANOVA, scatter plots, and correlations. It aimed to discover whether there was an effect for time (first and last lecture) or theme (lecture comprehension, lecturer English, and lecturer communicative ability) in students' perceptions of their lecture comprehension. The findings partly support H3.

A repeated measures ANOVA revealed no reliable results for time or theme. This finding is somewhat surprising: Given the increased range of variability in spoken ELF (due to non-standard deviations), the finding is contrary to common sense assumptions. Thus, the data were further explored. Using scatter plots showed marginal differences between the two time points, indicating a slight decline in the means for students' perceptions of lecture comprehension and of their perceptions of lecturer English. This difference suggests slight changes in students' perceptions over time although it does not indicate adjustment to lecturer English, thus lending partial support to H3. Nevertheless, the mean scores for lecture comprehension remained high at both points in time.

Additional exploration of the data through correlations revealed a relationship between students' perceived lecture comprehension and their perceptions of their lecturer's English. The positive association means that students' perceived their lecture comprehension to be better when they perceived their lecturer's English to be better. Moreover, the finding indicates the lecturer's English to be more important at Time 1 than at Time 2, which means that students adjusted to their lecturer's English or that their comprehension became less attached to their lecturer's English. This finding, which supports H3, should be further examined in future investigations, utilizing more sample times than two.

The discussion also explored possible reasons for apparent adjustments and non-adjustments to the lecturer's English, based on responses to Q34 and comments from Q35. Of the four aspects of lecturer English measured in Q34, the mean scores in the repeated measures suggest that students adjusted to fluency and intonation, but not to vocabulary or pronunciation of single words. Because these four variables are intertwined, the discussion examined the findings in relation to literature on adjusting to foreign accents in NS-NNS interaction since no ELF studies were found as well as to findings on lecture comprehension in EMI programs.

Findings from studies on L2 lecture comprehension reveal challenges with terminology and lecture content, regardless of whether the instructional language is the L1 or an L2. Studies that compare differences between L1 and L2 listening comprehension indicate little or no difference between proficient SLUs and those studying in their L1. In the present study, the findings suggest that some challenges with lecture content and terminology may be due to the inclusion of cross-disciplinary content that is atypical to

students' disciplinary fields. This finding aligns with previous studies on L2 lecture comprehension that show subjects from different fields to present more challenges than those in the primary field of study. In addition, a study at a Swedish university found lecture comprehension to be more challenging for first-year undergraduate students (of physics) in L2 English than in L1 Swedish. In short, the general picture suggests some challenges with key terminology and key concepts to be a common phenomenon in lecture comprehension, irrespective of L1 or advanced L2 proficiency. These studies also demonstrate that not all L1 speakers are proficient in lecture comprehension either.

In short, the moderate to high mean scores for perceived lecture comprehension at Time 1 and Time 2 in the present study suggests that students managed quite well and that they perceived their lecturer's English to be competent. The support for H3 suggests that proficient SLUs adjusted to some aspects of their lecturer's English, such as fluency and intonation, but not to others, such as vocabulary or pronunciation of single words. Other support for H3 also suggests that the lecturer's English becomes less important for lecture comprehension over time. These findings suggest future studies should examine these aspects of lecture comprehension more closely.

## 10 Conclusions and implications

At a time when universities outside the NS world are booming with courses and programs in English in order to attract international students, the need to ensure quality in teaching in English has taken on a new importance in a new dimension. Moreover, this new dimension particularly touches upon the job demands required of EMI lecturers. Thus, this study set out to explore what qualities should be considered in developing adequate criteria for assessing professional English in this international context. To achieve this, I focused on the micro level of language assessment through self-assessments of professional English by EMI university lecturers and examined the macro level through their students' perceptions of lecture comprehension. With this focus, I have been able to consider the insights provided by lecturers' own testimonies. The data proved to be fruitful as it shed light on criteria considered relevant for assessing EMI.

In this final chapter, I bring together the main themes that guided the study. In sections 10.1-.3, I summarize and discuss the three research questions analyzed in Chapters 7 through 9. By revisiting the research questions, I can make tentative claims about the results in relation to real-world assessment criteria for professional English in an international context. Then, in section 10.4, I synthesize the findings to address the overarching question in this study. Following this discussion, I present an evaluation of the study, and in the final section suggest further research.

### 10.1 What language ideologies guide EMI lecturers' self-assessments of their professional English?

This question (RQ1) was posed to investigate the ideological basis of English from the point of view of EMI lecturers. Using literature-driven ideologies, the aim was to detect what language ideologies guide the self-perceptions of these lecturers. In contrast to ELT professionals, who adhere closely to NS language ideologies (e.g. Jenkins 2007), it was expected that EMI lecturers might be less guided by this language ideology, and that their self-perceptions of professional English would be affected by the language ideologies informing them. A commonality between ELT professionals and EMI lecturers is that they have all been exposed to NS language ideologies in their education. However, the language targets and social contexts for the two groups differ (as discussed in 4.2), and most likely vary more among ELF users. For this reason, language ideologies must be studied in their social context. Because NS language ideologies uphold the belief in native-like language being correct and non-native language being deficient, it was hypothesized that EMI lecturers would have positive or negative perceptions of their professional English depending on how closely they perceived it to approximate the NS standard. The findings have implications for the development of language assessment criteria relevant to international contexts, as the following discussion illustrates.



In lecturer self-assessments, the analysis focused on language ideologies that were shared by the lecturers. It also examined whether language ideologies guided their positive and negative perceptions of their professional English. The analysis was approached from the level of the individual, providing a micro-analytic perspective.

Using the framework of language ideologies in Chapter 3, I imposed literature-driven ideologies on the data. With this method, I answered RQ1, where the results indicated that four widespread language ideologies guide EMI-lecturers' self-perceptions of their professional English. An important conclusion found context dependency to be a central tenet to the beliefs about language use and social behavior, which were grounded in the experiences of the individuals. For example, it is evident from the accounts that language ideologies related to standard language and NS language ideology bind the social actors to a particular nation-state, thereby constructing boundaries that define different groups, such as SLLs (assimilating to the nation-state) and SLUs (non-assimilating). What primarily emerged from the discussion is the application of different sociolinguistic appropriateness for different contexts, and the changing self-perceptions related to these contexts. This context-dependency has implications for language assessment criteria as it illustrates how closely language ideologies are intertwined with the local relevance of social context. This point is further illustrated in the discussion below.

Another important conclusion was the presence of multiple language ideologies, as well as conflicting ones. Contradictory views to the idealized NS were present. While NS accents are admired, they are also scorned. While natural fluency is appreciated, *perceived clarity or intelligibility* is key to some sort of 'neutral' accent (as opposed to native-likeness). EMI lecturers also view themselves as better communicators in ELF interaction than NSs. The role of the NS, however, was not dismissed. Rather than being viewed as a measure of near NS competence, it seemed to receive a new role – viewed as an idealized target that could help to reduce variation in ELF. This view included placing *accents perceived as 'neutral'* at the top of a hierarchy and 'strong' or 'incomprehensible' at the bottom, regardless of the accent being native or non-native. These findings suggest a comprehensibility goal of perceived clarity or intelligibility over native-likeness.

A similar view was also present in discourse related to standard language ideologies. Although linguistic correctness was claimed as important to communicating clearly, the view to adhering closely to this standard was perceived as a means to reduce broad variation in ELF, thereby enhancing ELF communication. Thus, the primary focus appears to be communicating efficiently, where linguistic correctness plays a secondary role. What the findings suggest is a focus on a comprehensibility goal as opposed to a correctness goal. These views to the NS contrast with findings from previous research on the ideological NS views upheld by ELT professionals (e.g. Jenkins 2007). Some beliefs that emerged from the present study include an emphasis on *clarity over accuracy*. For instance, *some NNS lexicogrammatical features*, such as non-native usage of articles and prepositions in phraseological sequences, are acceptable choices as long as they do not hinder comprehension of the main communicative goal. Another belief includes *self-correcting lexical items* important to comprehensibility (as opposed to linguistic form), such as word choice. A third belief is that their lexicogrammatical choices serve an important function: *to increase NNS fluency by drawing on their knowledge of the subject*

*matter and plurilingualism*, as demonstrated through the use of appropriate linguistic substitutions or mixing.

Conflicting views also emerged from discourse related to one-nation-one-language ideologies. For example, it was acceptable to sound Finnish and act Finnish in Finland – even when speaking English, but not outside Finland. This view illustrates the relevance of the local context in providing rules for sociolinguistic appropriateness. Moreover, the acquired local rules did not replace existing ones, but rather were added to them. This practice would most likely be perceived as inappropriate in an ENL target culture, and contrasts with NS language ideologies. In short, the beliefs point toward an understanding of local sociolinguistic behavior with acceptance toward variety in English that reflects *pluriculturalism* and *plurilingualism*.

The fourth ideology, English as a global language, contrasts with ideologies that support monolingua-culturalism. The English-as-a-global-language ideology is strongly linked to notions of language competence including plurilingua-cultural skills and knowledge, where there are no claims to NS ownership of the language or native-like proficiency. A belief central to this ideology is the principle of non-nativeness as binding. The view includes an acceptance of NNS features in ELF discourse and an ability to communicate successfully in ELF interaction. An example is substituting lexical items and grammatical structures from other languages as *a strategy to increase NNS fluency* as opposed to NS correctness that could result in lengthy pauses that signal dysfluency, which may be distracting to the listener. The communicative strategies are viewed as aids to maintaining flow in spoken ELF that support a comprehensibility goal. This goal and its related beliefs sharply contrast with beliefs linked to standard language and NS language ideologies, and consequently to the primary beliefs upheld by ELT professionals.

The connection between social context and language ideologies was also shown in the support for H1. The results showed a link between language ideologies and self-perceptions. The main finding revealed lecturers' self-perceptions of professional English changed depending on the language ideologies guiding their evaluations in a given context. For instance, in ENL contexts, the feeling of needing to conform to NS language ideologies resulted in negative perceptions of professional English that did not equate native-like English, viewing it as deficient. In contrast, in international contexts, the absence of native-likeness as a goal led to more positive views of professional English, seen as being 'good enough' or better than the 'average' NNS. What these differences also illustrate is context-dependency, with sociolinguistic appropriateness being tied to context. This context-dependency underscores the relevance of the social dimension in self-assessment, with both social context and agency as factors to be considered.

Overall, my investigation of EMI-lecturers' self-assessments of their professional English revealed their language ideologies to be based on the assumption of language being closely intertwined with context, which has local relevance. It also showed that individuals carry their histories with them, a point implicit and explicit in contradictory discourse related to the NS. This discourse included talking about NSs as though they were all the same. The view represents an idealized NS that serves as a model, one that was most likely introduced in language training at school and elsewhere. This model sometimes guided their views, which could result in self-contradiction. Alternatively,

other language ideologies that had been acquired during their complex histories with the English language were also present. The presence of alternative language ideologies, such as English as a global language, has implications for language assessment, as illustrated by EMI lecturers utilizing a different ideological base for self-assessing English. This point is further discussed in 10.4.

## **10.2 To what extent are the five CEFR scales a potential tool for assessing the professional English of EMI lecturers?**

This second research question was raised to examine lecturers' judgments of the generalizability of five CEFR scales to the assessment of their professional English for the purpose of lecturing in EMI. (For details on CEFR scale selection, see Chapters 1 and 5.) Including this aspect is important since meaningful testing should reflect the target situation. Thus, what the lecturers say in direct response to the CEFR criteria presented to them is extremely useful for pointing the way towards the central elements upon which relevant tests for this type of environment should be based. Including this research angle also adds pragmatic value to the present study. In other words, the aim was to investigate the extent to which a prescriptive, generic language assessment tool is suitable for measuring spoken professional English in an international context. For the investigation, a hypothesis was formed: the CEFR descriptor scales do not tap into the skills relevant to EMI lecturing.

To investigate the five CEFR scales, data were collected during semi-structured interviews. EMI-lecturers' self-assessments were based on their interpretations of the five scales intended for measuring the generic linguistic skills of SLLs – not the professional language of SLUs – as depicted by the prescriptive tones and deficit views in the scales (as presented in Chapter 5 and disclosed in the analysis in Chapter 8). The findings from the investigation support H2, indicating the scales to be inadequate for assessing the professional English of EMI lecturers.

What EMI lecturers considered as suitable criteria for assessing their professional English was not found in the CEFR scales. What they perceived as relevant to the assessment of their professional English is summarized here. For lecturing, important to *coherence and cohesion* is the ability to speak about conceptually more demanding matters with well-reasoned arguments using 'semantically dense' (cf. Fillmore 1979: 51) utterances. This skill entails the ability to use cohesive devices for linking utterances into coherent discourse as related to telling a story (logically). In a long contribution, some 'jumpiness' is perceived as normal. The view of jumpiness also includes a communicative function: it may serve as a repair strategy for retrieving logical progression, for example, through a dialogic act triggered by a question, which then either restores the line of thinking or takes it in a new direction. The cohesive devices were also perceived as being similar to those found in conversational discourse.

*Central to perceived fluency is (NNS) accessibility.* This criterion entails an ability to comprehend (NNS) fluency, where variability may be high. Some features of (NNS) accessibility include pausing when time is needed for message conceptualization as related

to formulating ideas and managing lexical gaps. Another criterion was *perceived intelligibility*. Important to its description is the concept of *a reasonable accent* (as defined by speech perception) that approximates what the intended audience could expect (whether ENL or not). In other words, an unreasonable accent includes speaking in a way that causes difficulty in comprehensibility. Closely related to this is the concept of *a 'correct' accent*, described as the ability to pronounce the sounds of English and to utilize nuclear stress to enhance communication. This view seems to support the notion of a lingua franca core as defined in Jenkins (2000). Another feature of perceived intelligibility includes *the ability to adjust or to accommodate others*, such as speaking slowly and clearly as opposed to fast-paced native-like speech. A fourth criterion is *acceptable NNS lexicogrammatical features*. Even at high levels of proficient English such features are present, and those that do not hinder comprehensibility or meaning are viewed as natural and acceptable.

For the purpose of assessing lecturing, the analysis also identified some relevant aspects to be missing from the CEFR scales. For example, the findings indicate important criteria to be *speaker credibility*, defined as including an ability to demonstrate in-depth knowledge of the subject matter, to handle questions well, and to speak with good grammar. Another important item is *making content accessible*, which includes an ability to present cognitively demanding content in a way that is easily understood. This criterion was further defined as pausing naturally and accommodatingly, as well as expressing ideas clearly through explanations and examples at an appropriate conceptual level. Having an ability to *engage the listener* was another criterion, defined as showing enthusiasm for and interest in the subject matter as well as speaking interactively. Correct use of *specialized vocabulary* was also considered important as well as knowledge of vocabulary from other fields when lecturing in a cross-disciplinary context.

For assessing LSP, the analysis also detected several *shortcomings* in the CEFR scales, some of which have been previously raised in the literature. Earlier research has pointed out the lack of descriptors for professional vocabulary and discourse features for professional communication. The present study targets generic prescriptive scales as the main drawback. In other words, the descriptors aim to assess general language abilities and lack strong performance measures rooted in real-world criteria that are necessary for assessing professional English in work-related tasks. This drawback results in scales that de-contextualize the 'learner', exclude situational appropriateness, omit communities of practice, minimize professional (i.e. specialized) vocabulary, and overlook indigenous assessment criteria. Thus, while a generic approach to language assessment may adequately evaluate general linguistic correctness, assessing LSP with such criteria creates a weak basis for adequately assessing the communicative character of work-related meanings and functions in a given real-world task.

The present study also found some *limitations* in the CEFR scales for assessing ELF. As already indicated in previous studies, the NS as the measure of language competence is a major drawback for the assessment of ELF. In the present study, lecturers' accounts suggest that the scales are interpreted as NS based, irrespective of whether they contain the term 'native speaker'. This term also represents a 'learner' goal of acculturating to a target ENL culture, a goal not shared by SLUs. As the yardstick of SLL attainment, the

CEFR emphasizes NS language ideologies and prescriptivism as the basis for assessment, which enforces both learner-related language ideologies and correctness as the measure of success. This measure evolves around a NS-NNS axis, where the aim for SLL is to assimilate to a given ENL target culture. These ideologies are prominent in the CEFR scales (as illustrated in Chapter 5), and contrast with its discourse related to plurilingualism and pluriculturalism. In brief, the emphasis on native-speakerism and correctness over the communicative character of the language fails to capture the nature of ELF spoken discourse.

All in all, the findings support H2, indicating that a generic tool, such as the CEFR, cannot adequately assess professional English (i.e. LSP). Nor can it assess English for use in an international context (i.e. ELF). Its prominent basis in NS language ideologies and prescriptivism contrasts with findings from the present study, which indicate that non-nativeness, variety, and plurilingua-culturalism are key factors for spoken professional English in an international context. These factors point toward the notion of *a competent international SLU* (see Table 4-2) as a concept against which to evaluate ELF usage. This notion contrasts with the idealized educated NS as an attainment target. While the target of an educated speaker seems plausible (see, for example, Mauranen 2012), the ELF literature casts doubt on the idealization aspect as being relevant to successful communication. This point crystallizes in the acceptability of (broad) variation in ELF interaction (see, for example, Hynninen 2013; Jenkins 2000; Mauranen 2012; and Ranta 2013). For LSP, the present findings indicate that the CEFR needs some sort of extra option. This option could include scales for assessing different genre, such as English for lecturing, advertising, or consulting.

### **10.3 What are EMI-students' perceptions of their lecturer's English during lectures, and do their perceptions change during the course?**

With EMI-students' perceptions of their lecturer's English, the focus shifted from the micro-analytic level to the macro. Whereas the analysis of language ideologies provided micro-level details from the perspective of the lecturers, the analysis of students' perceptions of English in lectures allowed me to examine the lecturer's English from the bigger picture, that is, at the classroom level. This angle was part of the triangulation strategies for the present study, providing a view to students' perceptions of their lecturer's spoken English. For the investigation, a hypothesis was formed. Because the capacity of humans to comprehend spoken language is highly flexible, it was hypothesized that students' perceptions of English in lectures would change over time as they adjusted to their lecturer's English. The hypothesis was investigated using repeated ANOVA, scatter plots, and correlations.

To examine change over time, a repeated ANOVA analysis was conducted. No reliable differences were found for time or theme. For this reason, the data were further explored using scatter plots and correlations.

Using scatter plots showed a slight decline in the marginal differences between the two points in time, which indicates some changes in students' perceptions. Moreover, the high mean scores at both points in time suggest that students comprehended the EMI lectures well. It also suggests a ceiling effect and that the measurement scale should perhaps have a broader range. The finding, however, confirms the results of students' self-assessments (presented in Chapter 6), which indicate the participants' English to be competent for managing EMI lectures.

Exploring the data with correlations revealed a positive association between students' perceptions of their lecture comprehension and of their lecturer's English. The relationship means that students' perceived their lecture comprehension to be better when they perceived their lecturer's English to be better. This finding suggests the lecturer's English to be more important for comprehension at Time 1 than at Time 2, indicating adjustment over the duration of a course. In other words, students' comprehension became less dependent on the lecturer's English over time, a finding that supports H3.

Reasons for apparent adjustments and non-adjustments to EMI-lecture comprehension were also explored. The mean score responses suggest that students adjusted to fluency and intonation, but not to vocabulary or pronunciation of single words. In addition, students' comments indicate some challenges with lecture content and terminology, and particularly to that which was atypical to their fields (for example in cross-disciplinary subjects). The findings support previous studies showing key terminology and concepts to present some challenges in lecture comprehension for students, whether using their L1 or being a proficient SLU. Moreover, students' comments confirmed that they are satisfied with their EMI lecturers and that they do not expect them to have 'perfect' English. This view to non-native English supports a differential view towards ELF as opposed to a deficit view. As discussed in Chapter 4, the deficit and differential views belong to opposing paradigms that are based on different language ideologies.

The conclusion here was that students adjusted to some aspects of their lecturer's English over the duration of a course, but not to all. This finding partly supports H3. Adjustments to fluency and intonation appear to have occurred more rapidly than to vocabulary use and pronunciation of single words. Non-adjustment to vocabulary use, however, aligns with findings from other studies on academic lecture comprehension, which show L1 and L2 speakers to have similar challenges. In addition, the high level of perceived confidence in EMI-lecturer English suggests that students perceived themselves and their lecturers to be proficient SLUs. In addition, the high mean scores for perceived lecture comprehension suggest that EMI students perceived their lecturer's English to be comprehensible. Students' comments on their lecturer's English also support a view of English being an instrument of communication in an international context, where 'perfect' English was not the focus, as discussed in the theoretical framework in 4.2.

## **10.4 What qualities of spoken professional English in an international context are relevant to developing assessment criteria? Implications**

The broad argument of this dissertation suggests a need to develop adequate language assessment criteria that is relevantly tuned to the target purpose and situation. This need is embedded in new sociolinguistic and demographic realities that are not addressed by current generic prescriptive assessment measures. To contribute to filling this need, the present study investigated perceptions of spoken professional English in an international context. The aim of this section is to synthesize the results into one composite answer to the overarching research question (presented in the header of this section).

All in all, the study of self-assessments of spoken professional English in relation to language ideologies has implications for the development of language assessment criteria. Results from the present study suggest that EMI lecturers have two basic views of their professional English: type A) when they compare themselves to ENL speakers, they perceive their English as faulty, and type B) when they think of themselves in their normal working environment (i.e. the target environment in this study), they view their English as working rather well. Moreover, these views changed when context changed, which illustrates context dependency. Similarly, their views changed when their chosen measure changed: They tended to point out their weaknesses when comparing themselves to ENL norms, and to notice their strengths when comparing themselves to another standard, such as the average Finnish speaker. These findings align with previous studies on language ideologies that show them to closely interface with social agency (see 3.2). The findings also imply different sociolinguistic appropriateness for different contexts based on different measures.

Findings from the student questionnaires also support the type B interpretation. Students' perceptions of their lecturer's English indicate a differential view to professional English in international contexts. In other words, students did not expect to hear 'perfect' English, and perceived their lecturer's English to be comprehensible. Their responses also indicated faster adaptation to fluency and intonation than to vocabulary and pronunciation of single words. Non-adjustment to vocabulary in lectures is not surprising: previous research into lecture comprehension also indicates key terminology and concepts to be a challenge for both L1 and proficient L2 speakers (e.g. Hellekjaer 2010).

The investigation thus indicates that certain language ideologies induce type A discourse and others type B. The discourse related to these is relevant to language assessment, since meaningful testing should reflect the target situation. My informants' responses to the CEFR scales are thus useful for pointing the way towards criteria upon which to base tests relevant to professional English in an international context. This view also supports the broad argument in the CEFR, which advocates the following: language competence for communication as related to the knowledge and skills necessary for the target situation and domain of use. The findings, however, do not support the narrow views in the descriptors (and elsewhere in the Framework) tied to native-speakerism and prescriptivism as relevant to an international context. Rather, what they imply is a focus

on the *comprehensibility* of spoken English in an international context. My findings include some criteria relevant to this context as summarized in Table 10-1.

**Table 10-1. Findings: Some relevant criteria for assessing spoken professional English in an international context.**

Criterion	Description
Accessible NNS fluency	<ul style="list-style-type: none"> <li>• Incorporates pausing when time is needed for message conceptualization, such as formulating ideas and managing lexical gaps, and appears natural in the context</li> <li>• Contains variability (in lexical, grammatical, and phonological items) whereby the speaker <ul style="list-style-type: none"> <li>• Demonstrates an ability to make him/herself understood</li> <li>• Utilizes strategies that enhance communication and prevent unnecessary gaps that could disrupt flow</li> </ul> </li> </ul>
Perceived intelligibility	<ul style="list-style-type: none"> <li>• Is defined by speech perception where an accent approximates what the intended audience could 'reasonably' expect</li> <li>• Involves some sort of 'neutrality' that contributes to perceived clarity and intelligibility (more than native-likeness), which includes speaking slowly and clearly</li> <li>• Includes core sounds essential to intelligibility across speakers with different L1s (e.g. as defined by Jenkins 2000)</li> </ul>
Accommodation skills	<ul style="list-style-type: none"> <li>• Entails an ability to adjust spoken language or to accommodate others in an effort to aid comprehensibility, such as adjusting the speaking pace, avoiding the use of idiomatic expressions, and using repetition and rephrasing.</li> </ul>
Plurilingua-cultural strategies	<ul style="list-style-type: none"> <li>• Comprise an ability to employ mixing, switching, and blending to enhance communication, including <ul style="list-style-type: none"> <li>○ Maintaining flow or NNS fluency</li> <li>○ Describing local items or concepts for which there are lexical gaps in English, thus utilizing a shared L2 repertoire from another language to overcome the situation</li> <li>○ Awareness of local and/or pluricultural politeness in international contexts</li> </ul> </li> </ul>
Lexicogrammatical appropriateness	<ul style="list-style-type: none"> <li>• Includes a focus on clarity (over accuracy) where variability does not hinder comprehensibility, such as articles and prepositions in phraseological sequences (e.g. discuss about, as the matter of fact, for a most part), regularization of verbs (e.g. choosed, taught) and nominal plurals (e.g. researches, equipments), and the speaker employs self-correction to clarify meaning (when necessary) important to comprehensibility</li> </ul>
Coherence & cohesion	<ul style="list-style-type: none"> <li>• Includes some 'jumpiness' in a long contribution, which may even serve as a repair strategy for retrieving logical progression, for example, through a dialogic act triggered by a question, which then restores the line of thinking or takes it in a new direction</li> <li>• In contributions that are conceptually more demanding, coherence includes the ability to speak with well-reasoned arguments using 'semantically dense' (cf. Fillmore 1979: 51) utterances</li> </ul>

The left-hand column outlines six criteria considered to be relevant to assessing spoken professional English in an international context. For each criterion, the description in the right-hand column drawn from the present study highlights tolerance toward variation in English: different accents where 'intelligibility' and 'NNS fluency' are more important than native-likeness, different lexicogrammatical usage where 'clarity' is more important than 'accuracy', and different pragmatics where plurilingua-culturalism and accommodation are valued over monolingua-culturalism and native-likeness.

For EMI lecturing, the present study found five criteria relevant to assessing English for this purpose, as shown in Table 10-2. Unlike the generic descriptors in the CEFR, these criteria are relevant to the professional context of use.



**Table 10-2 Findings – some relevant criteria for assessing lecturing.**

Criterion	Description
Content accessibility	<ul style="list-style-type: none"> <li>Involves an ability to present cognitively demanding content in an accessible manner, which includes pausing naturally and accommodatingly, as well as expressing ideas clearly through explanations and examples at an appropriate conceptual level</li> </ul>
Speaker credibility	<ul style="list-style-type: none"> <li>Includes an ability to demonstrate in-depth knowledge of the subject matter, to handle questions well, and to speak with good grammar</li> </ul>
Listener engagement	<ul style="list-style-type: none"> <li>Encompasses an ability to engage the listener through enthusiasm for and interest in the subject matter, as well as to speak interactively</li> </ul>
Field-specific terminology	<ul style="list-style-type: none"> <li>Includes the following abilities <ul style="list-style-type: none"> <li>Correct usage of field-specific terminology</li> <li>Adequate range of field-specific terminology</li> <li>Understanding of specialized vocabulary related to more than one specialized field, i.e. cross-disciplinary</li> </ul> </li> </ul>
Lecture coherence	<ul style="list-style-type: none"> <li>In a long contribution, some jumpiness is normal, and may even serve as a repair strategy for retrieving logical progression, for example, through a dialogic act triggered by a question, which then restores the line of thinking or takes it in a new direction</li> <li>In conceptually more demanding contributions, the user demonstrates the ability to speak with well-reasoned arguments using 'semantically dense' (cf. Fillmore) utterances</li> </ul>

What my analysis into the EMI-lecturer self-assessments indicates is that the competences relevant to assessing spoken professional English in an international context were not found in the five CEFR scales. As Table 10-1 shows, some criteria thought to be suitable for an international context include plurilingua-cultural strategies, an intelligible accent, accessible NNS fluency, lexicogrammatical 'clarity' (not accuracy), and accommodation strategies. Nor were the generic CEFR scales found to be appropriate for assessing lecturing. As shown in Table 10-2, some suitable criteria for this purpose include speaker credibility, listener engagement, content accessibility, lecture cohesion, and specialized terminology. These findings shed light on qualities of English relevant to assessing ELF and LSP. This kind of bottom-up approach is important for showing the ways in which speakers perceive their professional English for a given purpose and situation, and what they view as important for measuring a successful performance. This information can be used to inform the development of relevant criteria for assessing spoken professional English for international contexts.

An investigation into what constitutes a sufficient CEFR level for lecturing revealed differences in views between EMI lecturers and language and communication experts. While lecturers primarily view levels B1 and B2 as adequate, language experts consider C1 as the minimum level. The chief reason for this difference appears to be linked to the contrasting views to native-likeness. While EMI lecturers tended to choose CEFR descriptors that did not emphasize native-likeness as important to EMI lecturing, language and communication experts clearly did. This difference illustrates a contradiction in beliefs upheld by these two groups: many ELT professionals inherently equate 'good' English with native-like attainment (see Jenkins 2007) whereas non-ELT professionals, such as EMI lecturers, seem to equate it with a comprehensibility goal.

The EMI lecturer views in the present study present grounds for reconsidering what criteria might best measure spoken professional English in an international context. The findings imply that language assessment criteria for spoken professional English in an international context cannot be fundamentally derived from ineffective tools that do not conceptualize or confront the social context for which the assessment should be valid. The study indicates criteria pivoting around NS language ideologies to be quite distant from the reality of the professional needs of EMI lecturers. A top-down, policy-driven approach to language assessment based on prescriptive, generic measures does not represent the real-world needs of EMI lecturers. The findings point toward a comprehensibility goal over native-likeness as the primary measure of a successful performance. This finding supports previous studies that have reported NS language ideologies to be inadequate as a measure for ELF (Jenkins 2000, 2007). In order to assess L2 speakers adequately, measurement criteria must match the context of usage. As criteria for successful communication in an international context do not appear to be NS based, the findings also imply that ENL speakers would need to demonstrate that they have acquired adequate skills for using English in international contexts. In other words, they would not be exempted from an assessment of professional English for an international context based on NS abilities.

For language and communication experts, the results of this study represent a lay-speaker view: the primary subjects were lecturers of engineering, not linguists or ELT professionals. This aspect is evident in lecturers' comments, for instance, on vocabulary, accents, and grammar. One example is the view to 'common' words, which includes terms from their professional domain. While this may be the case for them, it is not the case for language or communication experts who have little or no knowledge of their specialty areas. This difference illustrates how vocabulary viewed by experts in one field (e.g. engineering) differs from those in another (e.g. linguistics), a point needing consideration particularly in the development of an LSP scale. Another example includes viewing one's own accent as normal while perceiving other accents as different or strong. In other words - others have accents, but I don't - a typical lay-speaker concept. A third example includes the idea that grammatical correctness means good communication, which is not true. Chomsky clearly illustrates this point in his famous nonsensical sentence that is perfectly grammatical, "Colorless, green ideas sleep furiously" (1957: 15). Some of the beliefs are self-contradictory, as the last example clearly illustrates. Moreover, some of these self-contradictory, lay-speaker views are present in the five CEFR scales (see, for example, Tables 5-1 and 5-4 in the present study). Such self-contradiction in assessment criteria raises questions about their validity, reliability, and fairness.

At the onset of this dissertation, I situated my research within a sociolinguistic framework. In this concluding chapter, it is natural to consider how the results relate to this perspective. In this dissertation, I have operated under the assumption that both micro- and macro-analytic views are important. This assumption is confirmed in the conclusion that lecturers' self-perceptions provide insight into assessing spoken professional English in an international context. I have also worked from the premise that language ideologies and context are closely intertwined and that personal histories will help to shape the views held. In other words, although sociocultural phenomena are socially constructed, there are

existing structures in place that will influence the perceptions of the actors. This premise bore out in my analysis that demonstrated a link between self-perceptions of language competence changing that depended on the language ideologies informing them in different contexts.

## **10.5 Evaluation of the study**

This study focused on the assessment of EMI lecturing because it involves the actual situation in which I work. The choice is thus based on a real-world need. Although my recommendations cover this genre best, many of the points are more general and thus highly relevant to the development of test batteries for spoken communication in EMI.

The study was largely based on the art of interpreting data gathered from semi-structured interviews and stimulated recall, using convenience sampling (Dörnyei 2007). It is hoped that this approach did not result in responsive sampling (Lincoln and Guba 1985) and that the methods produced credible data. For the interviews, the format and style appeared to encourage engagement with the topics as well as freedom to digress and develop emerging views that were important to the participants. The information collected during these meetings has something to say about lecturers' self-perceptions of their professional English as well as their ideas about what is relevant and important for assessing their skills.

The use of stimulated recall proved to be challenging, I think mainly due to my lack of experience in working with it. The protocol provided by Gass and Mackey (2000) proved beneficial, especially for setting up the sessions. However, as noted in Chapter 6, I did not closely follow the interview guide that I had prepared. This means that different questions were used to elicit data in each of the sessions. While this approach was not systematic, it was more natural and allowed the participants to focus on the parts of the videoed lecture that interested them. Because the topics of the lectures were unfamiliar to me, I did not attempt to edit the 45-minute videos. While I think this was a wise decision, it also meant that there were quite a few bits of lengthy silence during the recall sessions. Lecturers also commented on a wide range of items, including their visuals, voices, manners, mistakes, and aspects of the lecture topic itself. Although some fruitful data emerged, the effort put into collecting it seems greater than the fruits produced by it. To me, this suggests the stimulated recall was conducted with questions that were too general, and perhaps my attempts to remain in the background were too strong. It appears that the participants should have been prompted with more specific questions about their self-perceptions of their language competence and ability to teach in English. In short, this approach was used with rather limited success. Nevertheless, with the help of this method, it was possible to investigate lecturers' thinking in continuous connection with their actual performance in the classroom and to explore their perceptions of it. When the data collection takes place in the natural setting and in connection with the concrete events of the lecture, it can be assumed that these factors enable understanding of the investigated phenomenon more reliably. Adherence to a narrow window of time between the actual lecture event and the stimulated recall session also contributed to the reliability of the method.

As part of a longitudinal study, the repeated survey of students' perceptions worked well although it had some limitations. This instrument provided the means to investigate whether students' perceptions of their lecturer's English changed over the duration of a course (lasting either seven or fourteen weeks). Investigating perceptions over time gives not only good indicators of the reliability of the scales used, but also allows researchers to notice changes and explore the issues that most likely led to those changes. Primarily, what I learned from this instrument is that the measurement scale could have been wider and that more than two sample times would have been beneficial. In terms of the scale used, the range (e.g. 1-4) may have been too narrow. To prevent a ceiling effect, a broader range (e.g. 1-5) could have been employed. By adding an additional sample time, I could also have established more accurately the point at which perceived adjustment to the lecturer's English occurred during the course.

Another aspect related to sample times was the fluctuation of both student attendance and lecture content at initial and last lectures. Although the questionnaires were administered during regularly scheduled lectures, the paired response rate was 75 percent due to dropout rates during the first week of class as well as absences during the last lecture. Approaching students at the first and last lecture also had some disadvantages in terms of lecture content. For instance, in the first lecture, the focus is generally on giving an overview of the course including a brief introduction. In contrast, the last lecture can be a review of the main concepts or a complex topic that integrates many of the key concepts introduced throughout the course, making its complexity much higher than in the initial lecture. This difference can impact lecture comprehension that may have little or no relation to language proficiency. This point was commented by both the students and lecturers, and considered in the interpretation of the results.

The repeated survey was also limited to perceptions of EMI lecture comprehension. Ideally, students' L1 lecture comprehension would also have been checked. Out of thirty-six students, twelve had Finnish as an L1 and may have been attending comparable master's level courses in Finnish. Had I considered this initially, I could have pursued it as part of the present study. This element thus remains for a future study. Finally, the sample size was respectable but could have been larger. This aspect, however, requires extra effort and was beyond the scope of this work.

In terms of researcher position and impact, it is not easy to assess the degree to which my presence and the interpretation of it may have affected the responses. However, due to my additional role as mentor, I do believe that trust was established and that participants felt comfortable discussing their perceptions of lecturing in English. This aspect was evident in their sincere attempt to express their views, which were provided freely, purposefully, and insightfully. The interviews and stimulated recall were intended as conversations where the aim was dialogue providing "inter views" (cf. Kvale 1996). Nevertheless, my presence and planned questions could not have been invisible. My aim was to facilitate the discussion when needed, to ensure that some key questions were answered during the course of the interview, and to minimize my input in order to listen as much as possible. This approach resulted in messy data, as can be expected. It was also rich enough to provide sufficient examples for this study.

Any small-scale study has its limitations. The hope with small samples is that an in-depth understanding of the issues can be achieved as opposed to the breadth that larger samples can capture. The intention is that readers interested in the subject will bring their own experience to bear in transferring aspects of the findings and studying it in other contexts as opposed to claiming generalizability. This outcome is a consequence of a study partially grounded in an interpretative approach, where multiple voices are heard in a coherent piece, including that of the researcher. Moreover, researchers will inevitably bring some baggage to the task, whether related to methodological or substantive questions. I believe the effect of this research process includes a shifting of some perspectives, both on the part of the researcher and the participants. Thus, the next time would always be different, and a similar study would need to consider the sequence and spacing of the semi-structured interviews and stimulated recall. Alternatively, a focus group might provide a fuller picture of the competences EMI lecturers view as important for successfully lecturing in English and their positions in relation to the research questions. In retrospect, this study has value to me as a research practitioner with a stake in the issues. This dissertation is also an attempt to communicate its value to others.

## **10.6 The future**

In one sense the future is already here. With increased mobility, increased usage of English as a global lingua franca, and educational reforms designed to attract an international audience, institutions of higher education having undergone a transformation in order to provide EMI programs. In this postmodern world, language competence for SLUs of English is about the ability to maneuver between different types of Englishes, not how closely one can approximate a target ENL culture (Canagarajah 2006, Jenkins 2007). The current state of language assessment criteria, however, is lagging behind the recent developments in the use and users of English (Jenkins 2007). Although ELF is not fully described, sufficient research has been conducted to provide a basis to begin taking steps toward the development of suitable assessment criteria for SLU.

### **10.6.1 Taking a step forward**

Taking transformation as the theme, what can be concluded from this dissertation? Basically, that it is time to take a first step toward aligning language assessment criteria for this target purpose and situation. As a starting point, raising awareness about the abundant literature on ELF communication among the authorities who make decisions about the criteria for assessing EMI lecturers as well as among the assessment experts involved in the evaluation would be an important step. Raising awareness would, thus, be a first step toward implementing change in how EMI lecturers are evaluated and opening the door to developing suitable language assessment criteria for this purpose and situation. It is only through this realization that the focus of assessment could change towards a more realistic view of English in its current role as the world's primary lingua franca. Due

to the mainstream usage of ELF, research interest has mushroomed in recent years resulting in an abundance of ELF literature. Because of this devoted interest, there is now enough information available to begin to implement steps toward a more versatile means of assessing ELF than the NS communicative competence model. This step would help to reduce the mismatch between the ill-fitting NS language ideologies driving ELT assessment practices and the communicative objectives of SLUs.

With English as a global lingua franca, most likely language assessment will need to become broader and more diverse in order to develop assessments with local relevance. This means that language assessment will become less isolated from other fields and socially more responsive to local needs. This, in turn, means developing strong performance criteria for measuring professional language competence, rather than using weak performance measures heavily laden with linguistic criteria (as discussed in 4.1.2). For EMI lecturers, strong performance criteria aimed at measuring professional English highlights face validity. It also means that lecturers could focus on their teaching in an assessment without fear of making mistakes in English that would be judged as inferior to a NS, a point that is irrelevant to their teaching anyway – at least according to the findings in the present study where neither students nor lecturers expect native-like English. Any language assessment criteria that target the real-life purpose and situation should provide results that are more on target with the actual situation of target language use than those based on NS language ideologies, including the CEFR scales. Not only face validity of test results but also fairness to test takers requires changing the current language assessment criteria for English, which no longer represent the majority of English speakers in today's world.

To take an experimental first step towards developing suitable language assessment criteria for EMI lecturers, however, requires acceptance from the boards making decisions about what they see as a suitable target for EMI instructional purposes. Without this approval, there is little hope for innovative approaches to assessing EMI and little hope for specialized NNS lecturers with 'foreign accents' to qualify for tenured positions in EMI programs or to be hired as EMI lecturers in universities. As educational leaders, it is time for administrators in universities to show innovation once again. Taking such a step would require close cooperation with language professionals and support for developing assessment criteria suitable for measuring the spoken professional English of EMI lecturers.

### **10.6.2 Future research**

As the use of English in international programs in higher education has increasingly involved SLUs in recent years, the relevance of communication that is distant to NS norms has been increasingly questioned. If EMI lectures are in some sense the trend forward in international higher education, the significance of finding sound criteria for evaluating the professional English of these users increasingly grows. It is hoped that this study will facilitate design of appropriate criteria for this purpose and situation. Ultimately, such criteria will help to reduce the risk of an inappropriate language assessment caused by the

application of generic prescriptive assessment tools based on native speaker standards. Moreover, it seems likely that ELF usage for academic and professional purposes will remain prominent in the future as it is expected that students at the master's level will increasingly come from contexts where English is not the primary language. This situation makes determining standards for assessing EMI lecturing an important item on the research agenda of English language and communication.

For these reasons, I would encourage language professionals to work co-operatively towards developing a comprehensive research and development program that addresses establishing suitable assessment criteria for EMI lecturing. This co-operation could be conducted across universities and universities of applied sciences. In Europe, Pan-European cooperation would also be a great asset.

On assessing EMI lecturing, further investigation into what lecturers consider competent communication would provide additional insight into developing appropriate assessment criteria for this purpose and situation. While the present study provides some criteria as a starting point, further research would be essential for developing an assessment scale. Relevant to this is the question of how many scales would be sufficient and how the levels would be formulated.

In terms of other future research, a potential area to explore is the extent to which English language competence, both self-defined and as measured by standardized tests, is key to the ability to perform successfully in EMI lectures. From a lingua franca perspective, one question is the extent to which awareness of ELF is actually a feature of the professional experience.

On a related point, should all EMI lecturers be required to pass a certification test for teaching in English to a multilingua-cultural audience? Or does 'birthright' of nativeness ensure professional English for teaching in a multilingua-cultural context? These questions are interesting from the perspective of communication culture, which is part of the real world. Moreover, in the real world, communicative competence is oriented to performance competence, where all communicators, whether NSs or NNSs, are evaluated by what they do or do not do to achieve communicative success in a given situation for a given purpose. Studying the practices through which individuals accomplish successful performance in an indigenous assessment could illuminate how field-specific experts as cultural insiders determine what defines competent communication. For this reason, an interesting area of research would be to investigate the indigenous assessment criteria in different (academic) fields where there are a large number of EMI participants.

Alternatively, as a means to explore suitable language assessment criteria for academic professionals in EMI contexts, studies based on the ELFA corpus could be explored and used as an analytical baseline for compiling features common to academic ELF that are perceived as important to successful communication. Such a study could then serve as a basis for developing suitable assessment criteria for assessing academic ELF, which could then be piloted in another study.

## References

- Adank, Patti, Evans, Bronwen G, Stuart-Smith, Jane & Scott, Sophie K 2009. Comprehension of familiar and unfamiliar native accents under adverse listening conditions. *Journal of Experimental Psychology: Human Perception and Performance*, 35, 520-529.
- Adler, Max K 1977. *Pidgins, Creoles and Lingua Francas*, Hamburg, Buske Verlag.
- Ahvenainen, Tarmo 2005. *Problem-solving Mechanisms in Information Exchange Dialogues with English as a Lingua Franca*. Unpublished M.A. thesis, University of Jyväskylä. Available online: [http://thesis.jyu.fi/05/URN\\_NBN\\_fi\\_jyu-2005219.pdf](http://thesis.jyu.fi/05/URN_NBN_fi_jyu-2005219.pdf) [Accessed 10 June 2012].
- Airey, John. 2009. *Science, Language, and Literacy: Case Studies of Learning in Swedish University Physics*. Doctor of Philosophy in Physics with Specialization in Physics Education, Uppsala University.
- Airola, Anneli 2004. *Yritysten kielitaitotarpeet Pohjois-Karjalassa*. Joensuu: Pohjois-Karjalan ammattikorkeakoulu.
- Aitchison, Jean 1978. *Linguistics*. London: English Universities Press.
- Alderson, Charles, Clapham, Caroline & Wall, Diane 2005 [1995]. *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.
- Allen, Jim & Van Der Velden, Rolf 2005. The role of Self-Assessment in Measuring Skills. *REFLEX Working paper 2. Research Centre for Education and Labour Market*. Maastricht University, The Netherlands.
- Alves, Sonia Santos & Mendes, Lus 2006. Awareness and Practice of Plurilingualism and Intercomprehension in Europe. *Language and Intercultural Communication*, 6, 211-218.
- Anderson, Benedict 1983. *Imagined Communities*. London: Verso.
- Archibald, Alasdair, Cogo, Alessia & Jenkins, Jennifer 2011. *Latest Trends in ELF Research*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Ashton, Karen 2006. Can do self-assessment: investigating cross-language comparability in reading. *Research Notes*, 11-14.
- Bachman, Lyle 1990. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bachman, Lyle 2004. *Statistical Analysis for Language Assessment*. Cambridge: Cambridge University Press.
- Bachman, Lyle & Palmer, Adrian 1996. *Language Testing in Practice*. Oxford: Oxford University Press.
- Baumgardner, Robert J & Brown, Kimberley 2003. World Englishes: ethics and pedagogy. *World Englishes*, 22, 245-251.
- Becker, Gary S. 1964. *Human capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: NBER.
- Belcher, Diane & Lukkarila, Lauren 2011. Identity in the ESP Context: Putting the Learner Front and Center in Needs Analysis. In: Belcher, Diane, Johns, Ann M & Paltridge, Brian (eds.) *New Directions in English for Specific Purposes Research*. Ann Arbor: Michigan University Press.
- Benesch, Sarah 1996. Needs analysis and curriculum development in EAP: An example of a critical approach. *TESOL Quarterly*, 30, 723-738.
- Benesch, Sarah 1999. Rights Analysis: Studying power relations in an academic setting. *English for Specific Purposes*, 18, 313-327.



- Benesch, Sarah 2001. *Critical English for academic purposes: Theory, politics and practice*. Mahwah, NJ: Lawrence Erlbaum.
- Bent, Tessa & Bradlow, Ann R 2003. The interlanguage speech intelligibility benefit. *Journal of Acoustical Society of America*, 114, 1600-1610.
- Berggren, Ola 1973. *Kieltenopetuksen uudelleenjärjestäminen Lappeenrannan teknillisessä korkeakoulussa [The reorganization of foreign language teaching at Lappeenranta University of Technology]*. Jyväskylä: Jyväskylän yliopisto.
- Berggren, Ola 1975. *Teollisuuden palveluksessa toimivien insinöörien kielitaidon tarve ja insinöörökunnan käsitykset teknillisissä korkeakouluissa annettavan kieltenopetuksen sisällöstä*. Lappeenranta: Kieltenopetuksen keskuksen julkaisuja / Lappeenrannan teknillinen korkeakoulu.
- Berggren, Ola 1979. *Teollisen know-how'n vienti ja projektihenkilökunnan kielikoulutus: Jari-projektin Portugalin opetuksen tulokset ja niistä vedettävät johto päätökset [The export of engineering know-how and the language training of the project personnel: the results and conclusions of the teaching of Portuguese to the participants in the Jari-project]*. Lappeenranta: Kielikeskuksen julkaisuja / Lappeenrannan teknillinen korkeakoulu.
- Berry, V. E. 2004. *A study of the interaction between individual personality differences and oral performance test facts*. King's College: London.
- Biber, Douglas 2006. *University Language: A Corpus-based Study of Spoken and Written Registers*. Amsterdam: John Benjamins Publishing.
- Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan & Finegan, Edward 1999. *Longman Grammar of Spoken and Written Language*. London: Longman.
- Björkman, Beyza 2010. *Spoken lingua franca English at a Swedish technical university*. Doctoral dissertation, Department of English, Stockholm University.
- Blackledge, Adrian 2000. Monolingual ideologies in multilingual states: Language, hegemony and social justice in Western liberal democracies. *Estudios de Sociolingüística*, 1, 25-45.
- Bley-Vroman, Robert 1989. What is the logical problem of foreign language learning? In: Gass, Susan & Schachter, Jacquelyn (eds.) *Linguistic perspectives on second language acquisition*. Cambridge: Cambridge University Press.
- Blommaert, Jan 2006. Language Ideology. In: Brown, Keith (ed.) *Encyclopedia of Language and Linguistics*. Second edition. Oxford: Elsevier.
- Blommaert, Jan 2007. On scope and depth in linguistic ethnography. *Journal of Sociolinguistics*, 11, 682-688.
- Blommaert, Jan 1999. The debate is open. In: Blommaert, Jan (ed.) *Language Ideological Debates*. Berlin: Mouton de Gruyter.
- Blommaert, Jan & Verschueren, Jef 1998. The role of language in European nationalist ideologies. . In: Schieffelin, Bambi B., Woolard, Kathryn A. & Kroskrity, Paul V. (eds.) *Language Ideologies: Theory and Practice*. New York: Oxford University Press.
- Boswood, Tim 1999. Redefining the professional in international professional communication. In: Lovitt, Carl R & Goswami, Dixie (eds.) *Exploring the Rhetoric of International Professional Communication: An Agenda for Teachers and Researchers*. Amitville: NY Baywood Publishing Company Inc.
- Bourdieu, Pierre 1991. *Language and Symbolic Power*. Cambridge: Polity Press.
- Bradley, Henry 1904. *The Making of English*. New York: The MacMillan Company.
- Bradlow, Ann R & Bent, Tessa 2008. Perceptual adaptation to non-native speech. *Cognition*, 106, 707-729.
- Braine, George 1999. *Non-native Educators in English Language Teaching*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

- Braine, George 2005. A history of research on non-native speaker English teachers. In: Llurda, Enric (ed.) *Non-native Language Teachers: Perceptions, Challenges and Contributions to the Profession*. New York: Springer.
- Briggs, Charles L 1996. Disorderly Dialogues in Ritual Impositions of Order: The Role of Pragmatics in Warao Dispute Mediation. In: Briggs, Charles L (ed.) *Disorderly Discours: Narrative, Conflict, and Inequality*. New York: Oxford University Press.
- Brown, James D 1990. Evaluating one's abilities: Shortcuts and stumbling blocks on the road to self-knowledge. *Journal of Experimental Social Psychology*, 26, 149-167.
- Bryman, Alan 2004. *Social Research Methods*. Oxford: Oxford University Press.
- Calderhead, James 1981. A psychological approach to research on teachers' classroom decision making. *British Educational Research Journal*, 7, 51-57.
- Calderheard, James 1981. Stimulated recall: A method for research on teaching. *British Journal of Educational Psychology*, 51, 211-217.
- Cameron, Deborah 1995. *Verbal Hygiene*. London: Routledge.
- Camiciottoli, Belinda Crowford 2005. Adjusting a business lecture for an international audience: a case study. *English for Specific Purposes*, 24, 183-199.
- Canagarajah, A. Suresh 1999. *Resisting Linguistic Imperialism in English Teaching*. Oxford: Oxford University Press.
- Canagarajah, A. Suresh 2002. *Critical Academic Writing and Multilingual Students*. Ann Arbor: University of Michigan Press.
- Canagarajah, A. Suresh 2005. Changing communicative priorities, shifting pedagogical priorities, revised assessment objectives. *Symposium paper presened at the 14th World Congress of Applied Linguistics*. Madison, WI.
- Canagarajah, A. Suresh 2006. Changing communicative needs, revised assessment objectives: Testing English as an international language. *Language Assessment Quarterly*, 3, 229-242.
- Canale, Michael 1983. On some dimensions of language proficiency. In: Oller, J.W (ed.) *Issues in Language Testing Research*. Rowley, MA: Newbury House.
- Canale, Michael 1984. A communicative approach to language proficiency assessment in a minority setting. In: Rivera, Charlene (ed.) *Communicative Competence Approaches to Language Proficiency Assessment: Research and Application*. Clevedon, England: Multilingual Matters.
- Canale, Michael & Swain, Merrill 1980. Theoretical bases of communicative approaches to language teaching and testing. *Applied Linguistics*, 1, 1-47.
- Carroll, John B 1961 [1972]. Fundamental considerations in testing for English language proficiency in foreign students. In: B, Allen Harold & N, Campbell Russell (eds.) *Teaching English as a Second Language: A Book of Readings*. Second edition. New York: McGraw-Hill.
- Carroll, John B 1968. The psychology of language testing. In: Davies, Alan (ed.) *Language Testing Symposium: A Psycholinguistic Perspective*. London: Oxford University Press.
- Carter, Ronald & Mccarthy, Michael 2006. *Cambridge Grammar of English: A Comprehensive Guide: Spoken and Written English: Grammar and Useage*. Cambridge: Cambridge University Press.
- Chambers, Jack K & Trudgill, Peter 1998. *Dialectology*. Cambridge: Cambridge University Press.
- Channell, Joanna 1994. *Vague Language*. Oxford: Oxford University Press.
- Chomsky, Noam 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam 1965. *Aspects of the Theory of Syntax*. Boston: MA, MIT Press.

- Clapham, Caroline. 1994. *The Effect of Background Knowledge on EAP Reading Test Performance*. Doctor of Philosophy, Lancaster University.
- Clapham, Caroline 1996. *The Development of IELTS: A Study of the Effect of Background Knowledge on Reading Comprehension*. Cambridge: Cambridge University Press.
- Clark, Constance M & Garrett, Merrill F 2004. Rapid adaptation of foreign-accented English. *Journal of Acoustical Society of America*, 116, 3647-3658.
- Clark, John L D 1972. *Foreign Language Testing: Theory and Practice*. Philadelphia, PA, Center for Curriculum Development, Inc.
- Cogo, Alessia 2009. Accommodating difference in ELF conversations: A study of pragmatic strategies. In: Mauranen, Anna & Ranta, Elina (eds.) *English as a Lingua Franca: Studies and Findings*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Cogo, Alessia & Dewey, Martin 2006. Efficiency in ELF communication: from pragmatic motives to lexico-grammatical innovation. *Nordic Journal of English Studies*, 5, 59-93.
- Cohen, Andrew D & Hosenfeld, Carol 1981. Some uses of mentalistic data in second language research. *Language Learning*, 31, 285-313.
- Cohen, Jacob 1988. *Statistical Power Analysis for the Behavioral Sciences*. New Jersey: Lawrence Erlbaum.
- Cohen, Louis, Manion, Lawrence & Morrison, Keith 2007. *Research Methods in Education*. Abingdon, Oxon: Routledge.
- Collins, James 1999. The Ebonics controversy in context: Literacies, subjectivities and language ideologies in the United States. In: Blommaert, Jan (ed.) *Language Ideological Debates*. Berlin: Mouton de Gruyter.
- Commission, European 2003. *Promoting Language Learning and Linguistic Diversity: An Action Plan 2004-2006*. Brussels: COM (2003) 449 final.
- Commission, European 2005a. *A New Framework strategy for Multilingualism*. Brussels: COM (2005) 596 final.
- Commission, European 2007. Focus on the Structure of Higher Education in Europe: 2006/2007 National Trends in the Bologna Process. *Directorate-General for Education and Culture*. Brussels: Eurydice.
- Commission, European 2009. *The Bologna Process: towards the European Higher Education Area*. [http://ec.europa.eu/education/higher-education/doc1290\\_en.htm](http://ec.europa.eu/education/higher-education/doc1290_en.htm). [Accessed 04 April 2012].
- Commission, European Communities 2005b. *A New Framework Strategy for Multilingualism*. Brussels: European Commission.
- Cook, Vivian 1997. Monolingual bias in second language acquisition research. *Revista Canaria de Estudios Ingleses*, 35-50.
- Cook, Vivian 1999. Going beyond the native speaker in language teaching. *TESOL Quarterly*, 33, 185-209.
- Cook, Vivian 2001. *Second Language Learning and Language Teaching*. London: Arnold.
- Corder, S. Pit 1981. *Error Analysis and Interlanguage*. Oxford: Oxford University Press.
- Coulmas, Florian (ed.) 1981. *A Festschrift for native speaker*. The Hague: Mouton.
- Cox, C. Brian 1989. English for Ages 5 to 16. In: *Des, Department of Education and Science* (ed.). London: Her Majesty's Stationery Office.
- Creswell, John 2003. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks, CA: Sage Publications.
- Crystal, David 1985. *A Dictionary of Linguistic and Phonetics*. Second edition. Oxford: Basil Blackwell.
- Crystal, David 2003. *English as a Global Language*. Cambridge: Cambridge University Press.

- Davies, Alan 1968. *Language Testing Symposium. A Psycholinguistic Perspective*. London: Oxford University Press.
- Davies, Alan 1991. *The Native Speaker in Applied Linguistics*. Edinburgh: Edinburgh University Press.
- Davies, Alan 2003. *The Native Speaker: Myth and Reality*. Clevedon: Multilingual Matters
- De Graaff, Erik, Demlova, Marie, Kuru, Selahattin & Peltola, Hannu 2007. Innovative Learning and Teaching Methods. In: Borri, Claudio & Maffioli, Francesco (eds.) *Erasmus Thematic Network*. Firenze: Firenze University Press.
- De Saussure, Ferdinand 1915. *Lecture Series on a Course in General Linguistics*. Geneva: University of Geneva.
- Deardorff, Darla K 2006. Identification and Assessment of intercultural Competence as a Student Outcome of Internationalization. *Journal of Studies in International Education*, 10, 241-266.
- Dewey, Martin 2009. English as a lingua franca: heightened variability and theoretical implications. In: Mauranen, Anna & Ranta, Elina (eds.) *English as a Lingua Franca: Studies and Findings*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Dicker, Susan J. 1996. *Language in American: A Pluralist View*. Clevedon: Multilingual Matters.
- Dörnyei, Zoltan 2003a. Attitudes, orientations, and motivations in language learning: Advances in theory, research, and applications. *Language Learning*, 53, 3-32.
- Dörnyei, Zoltan 2003b. *Questionnaires in Second Language Research: Construction, Administration, and Processing*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.
- Dörnyei, Zoltan 2007. *Research Methods in Applied Linguistics*. Oxford: Oxford University Press.
- Dörnyei, Zoltan & Kormos, Judit 1998. Problem-solving mechanisms in L2 communication: A psycholinguistic perspective. *Studies in Second Language Acquisition*, 20, 349-386.
- Douglas, Dan 2000. *Assessing Languages for Specific Purposes*. Cambridge: Cambridge University Press.
- Douglas, Dan & Myers, Ron 2000. Assessing the communication skills of veterinary students: Whose criteria? In: Kunnan, Antony John (ed.) *Fairness and Validation in Language Assessment: Selected Papers from the 19th language research colloquium*. Cambridge, UK: University of Cambridge Local Examinations Syndicate.
- Dudley-Evans, Tony & St. John, Maggie Joe 1998. *Developments in English for specific purposes*. London: Cambridge University Press.
- Dunning, David 2005. *Self-insight: Roadblocks and Detours on the Path to Knowing Thyself*. New York: Psychology Press.
- Dunning, David, Heath, Chip & Suls, Jerry M 2005. Flawed self-assessment: Implications for health, education, and the workplace. *Psychology Science in the Public Interest*, 5, 69-106.
- Ehrenreich, Susanne (ed.) 2009. *English as a lingua franca in multinational corporations - Exploring business communities of practice*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Ehrenreich, Susanne (ed.) 2011. The Dynamics of English as a Lingua Franca in International Business: A Language Contact Perspective. In: Alasdair Archibald, Alessia Coco, & Jennifer Jenkins (ed.) *Latest Trends in ELF Research*. Newcastle upon Tyne: Cambridge Scholars Publishing
- Elder, Catherine 1993. How do subject specialists construe language proficiency? *Language Testing*, 10, 235-254.

- Elder, Catherine 1994. Performance testing as benchmark for LOTE teacher education. *Melbourne Papers in Language Testing*, 3, 1-25.
- Ellis, Rod 2003. *Task-based language learning and teaching*. Oxford: England, Oxford University Press.
- Erickson, Frederick & Mohatt, Gerald 1977. *The Social Organization of Participation Structure in Two Classrooms of Indian Students*. Ottawa, Canada, Department of Indian Affairs and Northern Development.
- Eu 1995. *White Paper on Education and Training: Teaching and Learning*. Available online: [http://europa.eu/documents/comm/white\\_papers/pdf/com95\\_590\\_en.pdf](http://europa.eu/documents/comm/white_papers/pdf/com95_590_en.pdf) [Accessed 17 April 2012].
- Europe, Council of 2001. *Common European Framework of Reference for Languages: Learning, Teaching, and Assessment*. Cambridge: Cambridge University Press.
- Europe, Council of 2005. *Survey on the use of the Common European Framework of Reference for Languages: Synthesis of results*. Available online: <http://www.coe.int/t/dg4/linguistic/Source/Surveyresults.pdf> [Accessed 5 April 2012].
- Europe, Council of 2007. *From Linguistic Diversity to Plurilingual Education: Guide for the Development of Language Education Policies in Europe*. In: Division, Language Policy (ed.) Original 2002, revised 2007 ed. Strasbourg, Council of Europe: [http://www.coe.int/t/dg4/linguistic/Sorce/Guide\\_Main\\_Beacco2007\\_EN.doc](http://www.coe.int/t/dg4/linguistic/Sorce/Guide_Main_Beacco2007_EN.doc) [Accessed 17 April 2012].
- Eurostat 2012. *Foreign Language Learning Statistics*. European Commission. Online: [http://epp.eurostat.ec.europa.eu/statistics\\_explained/index.php/Foreign\\_language\\_learning\\_statistics](http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Foreign_language_learning_statistics) [Accessed 18 April 2012].
- Eva, Kevin W & Regehr, Flenn 2005. Self-assessment in the Health Professions: A Reformulation and Research Agenda. *Academic Medicine*, 80, Supplement, S46-S54.
- Faerch, Claus & Kasper, Gabrielle 1987. From product to process – Introspective methods in second language research. In: Faerch, Claus & Kasper, Gabrielle (eds.) *Introspection in Second Language Research*. Clevedon: Multilingual Matters.
- Fairclough, Norman 1989 [2001]. *Language and Power*. Harlow, England: Pearson Education Limited.
- Fairclough, Norman 1995. *Critical Discourse Analysis: The Critical Study of Language*. London and New York, Longman.
- Field, John 2003. *Psycholinguistics: A Resource Book for Students*. London: Routledge.
- Field, John 2005. Intelligibility and the listener: the role of 'lexical stress'. *TESOL Quarterly*, 39, 339-424.
- Fillmore, Charles J. 1979. On Fluency. In: Fillmore, Charles J., Kempler, Daniel & Wang, William S.-Y. (eds.) *Individual Differences in Language Ability and Language Behavior*. New York: Academic Press.
- Fillmore, Charles J. 1979 [2000]. On fluency. In: Rigggenbach, Heidi (ed.) *Perspectives on Fluency*. Ann Arbor: University of Michigan Press.
- Firth, Alan 1990. 'Lingua franca' negotiations: Towards an interactional approach. *World Englishes*, 9, 269-280.
- Firth, Alan 1996. The discursive accomplishment of normality: On 'lingua franca' English and conversation analysis. *Journal of Pragmatics*, 26, 237-259.
- Firth, Alan 2009. Doing not being a foreign language learner: English as a lingua franca in the workplace and (some) implications for SLA. *International Review of Applied Linguistics in Language Teaching*, 43, 127-156.
- Firth, Alan & Wagner, Johannes 1997. On discourse, communication, and (some) fundamental concepts in SLA research. *Modern Language Journal* 81, 285-300.

- Flege, James & Fletcher, Kathryn L 1992. Talker and listener effects on degree of perceived foreign accent. *Journal of Acoustical Society of America*, 91, 370-389.
- Flowerdew, John (ed.) 1995. *Academic Listening: Research Perspectives*. Cambridge: Cambridge University Press.
- Flowerdew, John & Miller, Lindsay 1992. Student perception, problems and strategies in SL comprehension. *Journal of Language Learning and Research in South Asia*, 23, 60-80.
- Fontana, Andrea & Frey, James H 2005. The interview. In: Denzin, Norman K & Lincoln, Yvonna S (eds.) *The Sage Handbook of Qualitative Research*. Thousand Oaks, California: Sage.
- Foster, Pauline & Skehan, Peter 1996. The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18, 299-323.
- Friedman, Thomas L 2005. *The World is Flat: The Globalized World in the Twenty-first Century*. London: Penguin Books.
- Fritzpatrick, Robert & Morrison, Edward J 1971. Performance and product evaluation. In: Thorndike, Robert L (ed.) *Educational Measurement*. Second edition. Washington, DC: American Council on Education.
- Gal, Susan 1993. Diversity and contestation in linguistic ideologies: German speakers in Hungary. *Language in Society*, 22, 337-359.
- Gal, Susan 1998. Multiplicity and Contention among Language Ideologies: A commentary. In: Schieffelin, Bambi B., Woolard, Kathryn A. & Kroskrity, Paul V. (eds.) *Language Ideologies: Practice and Theory*. Oxford: Oxford University Press.
- Gal, Susan & Irvine, Judith T. 1995. The boundaries of languages and disciplines: How ideologies construct difference. *Social Research*, 62, 967-1001.
- Gal, Susan & Woolard, Kathryn A 1995. Constructing languages and publics: Authority and representation. *Pragmatics* 5, 129-139.
- Gallois, Cindy, Ogay, Tania & Howard, Giles 2005. Communication accommodation theory: A look back and a look ahead. In: Gudykunst, William B (ed.) *Theorizing About Intercultural Communication*. London: Sage.
- Garfinkel, Harold 1967. *Studies in Ethnomethodology*. Englewood Cliffs, NJ: Prentice-Hall.
- Gass, Susan M. & Mackey, Alison 2000. *Stimulated Recall Methodology in Second Language Research*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Geertz, Clifford 1983. *Local Knowledge: Further Essays in Interpretive Anthropology*. New York: Basic Books.
- Giddens, Anthony 1984. *The Constitution of Society: Outline of the Theory of Structuration*. Cambridge: Cambridge University Press.
- Gilbert, Nigel & Mulkay, Michael 2003. *Opening Pandora's Box*. Cambridge: Cambridge University Press.
- Giles, Howard 1973. Accent mobility: A model and some data. *Anthropological Linguistics*, 33, 27-42.
- Giles, Howard 1977. *Language, Ethnicity, and Intergroup Relations*. New York: Academic Press.
- Giles, Howard & Coupland, Nikolas 1991. *Language: Context and Consequences*. Milton Keynes [England]: Open University Press.
- Gliner, Jeffrey A 1994. Reviewing qualitative research: Proposed criteria for fairness and rigor. *The Occupational Therapy Journal of Research* 14, 78-90.
- Goodwin, Charles & Goodwin, Marjorie H 1996. Crafting vision: Perception as professional practice. In: Gunnarsson, Britt-Louise, Linnel, Per & Nordberg, Bengt (eds.) *The Construction of Professional Discourse*. London: Longman.

- Graddol, David 2006. *English Next*. London: The British Council.
- Graham, Gibbs 1988. *Learning by Doing: A guide to Teaching and Learning Methods, Oxford Centre for Staff and Learning Development*. Oxford Polytechnic. London: Further Education Unit.
- Gramsci, Antonio 1971. *Selections from the Prison Notebooks*. New York: International.
- Griffiths, Roger 1990. Speech rate and NNS comprehension: A preliminary study in time-benefit analysis. *Language Learning*, 40, 311-336.
- Griffiths, Roger 1992. Speech rate and listening comprehension: Further evidence of the relationship. *TESOL Quarterly*, 26, 385-391.
- Haberland, Hartmut 2009. English - The language of globalism. *Rask Internationalt tidsskrift for sprog og kommunikation*, 30, 17-45.
- Hakala, Henrik. 2007. "Almost as annoying as the Yank; better accent, though." - *Attitudes and conceptions of Finnish students toward accents of English*. M.A. thesis. University of Helsinki.
- Hample, Dale 1984. On the use of self-reports. *Journal of the American Forensic Association*, 20, 140-153.
- Harris, Mary. 2002. Anti-globalisation and ELT. *IATEFL Issues*, 6-7.
- Hellekjaer, Glenn Ole 2010. Language matters: Assessing lecture comprehension in Norwegian English-medium higher education. In: Dalton-Puffer, Christiane, Nikula, Tarja & Smit, Ute (eds.) *Language Use and Language Learning in CLIL Classrooms*. Amsterdam: John Benjamins.
- Heller, Monica 1999. Language choice, social institutions and symbolic domination. *Language in Society*, 24, 373-405.
- Hine, Alison 2000. Mirroring effective education through mentoring, metacognition and self reflection. *Australian Association for Research in Education Conference*. Available online: <http://www.aare.edu.au/00pap/hin00017.htm> [Accessed 31 July 2012].
- Holborrow, Marnie 1999. *The Politics of English: A Marxist View of Language*. London: Sage.
- Holliday, Adrian 2005. *The Struggle to Teach English as an International Language*. Oxford: Oxford University Press.
- Horppu, Ritva & Lehtonen, Tuula 2003. *Monipuolinen kielitaito? Selvitys Helsingin yliopistosta valmistuneiden kielioopinnoista*. Helsinki: Helsingin Yliopisto Opintoasiainosaston julkaisu 22.
- House, Julianne 1999. Misunderstanding in intercultural communication: Interactions in English as a lingua franca and the myth of mutual intelligibility. In: Gnutzmann, Claus (ed.) *Teaching and Learning English as a Global Language: Native and Non-native Perspectives*. Tübingen: Stauffenburg.
- House, Julianne. 2001. A stateless language that Europe must embrace. *The Guardian*.
- House, Julianne 2003. English as a lingua franca: A threat to multilingualism? *Journal of Sociolinguistics*, 7, 556-578.
- Huang, Jinyan 2004. Voices from Chinese students: Professor's use of English affects academic listening. *College Student Journal*, 38, 212-223.
- Huckin, Thomas & Olsen, Leslie 1990. Point-driven understanding in engineering lecture comprehension. *English for Specific Purposes*, 9, 33-47.
- Hudson, Thom 2005. Trends in assessment scales and criterion-referenced language assessment. *Annual Review of Applied Linguistics* 25, 205-227.
- Hughes, Rebecca 2008. Internationalization in higher education and language policy: questions of quality and equity. *Higher Education Management and Policy*, 20, 111-128.

- Huhta, Marjatta 1997. *The Dynamics of Language Training - from an Element of Cost in to an Investment in Communication*. Helsinki: Opetushallitus.
- Huhta, Marjatta 1999. *Language/Communication Skills in Industry and Business*. Helsinki: Opetushallitus.
- Huhta, Marjatta. 2010. *Language and Communication for Professional Purposes - Needs Analysis Methods in Industry and Business and their Yield to Stakeholders*. Doctor of Science in Technology, Helsinki University of Technology.
- Hülmbauer, Cornelia 2009. "We don't take the right way. We just take the way we think you will understand" - the shifting relationship between correctness and effectiveness in ELF. In: Mauranen, Anna & Ranta, Elina (eds.) *English as a lingua franca: Studies and Findings*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Hüttner, Julia 2009. *Fluent Speakers - Fluent Interactions: On the Creation of (Co)-Fluency in English as a Lingua Franca*. In: Mauranen, Anna & Ranta, Elina (eds.) *English as a Lingua Franca: Studies and Findings*. Newcastle upon Tyne: Cambridge Scholars.
- Hymes, Dell 1967. Models of the interection of langauge and social setting. *Journal of Social Issues*, 23, 8-38.
- Hymes, Dell 1972a. Models of interaction and social life. In: Gumprez, John & Hymes, Dell (eds.) *Directions in Sociolinguistics: The Ethnography of Communication*. New York: Holt, Rinehart and Winston.
- Hymes, Dell 1972b. On communicative competence. In: Pride, J.B. & Holmes, J. (eds.) *Sociolinguistics*. Harmondsworth:: Penguin.
- Hynninen, Niina. 2006. *Cultural Discourses in CEF: How Do They Relate to ELF?* Master of Arts thesis, University of Helsinki.
- Hynninen, Niina 2010. "We try to to to speak all the time in easy sentences" - Student conceptions of ELF interaction. *Helsinki English Studies*, 6, 29-43.
- Hynninen, Niina 2011. The practice of 'mediation' in English as a lingua franca interaction. *Journal of Pragmatics*, 965-977.
- Hynninen, Niina. 2013. *Language Regulation in English as a Lingua Franca: Exploring Langauge-Regulatory Practices in Academic Spoken Discourse*. Doctor of philosophy, University of Helsinki.
- Irvine, Judith 1989. When talk isn't cheap: Language and political economy. *American Ethnologist*, 16, 248-267.
- Irvine, Judith & Gal, Susan 2000. Language ideology and linguistic differentiation. In: Kroskrity, Paul V. (ed.) *Regimes of Language: Ideologies, Politics, and Identities*. Santa Fe, New Mexico: School of American Research Press.
- Jacoby, Sally 1998. *Science as Performance: Socializing Scientific Discourse through the Conference Talk Rehearsal*. Doctor of Philosophy Doctoral dissertation, University of California, Los Angeles.
- Jacoby, Sally & Gonzales, Patrick 1991. The constitution of expert-novice in scientific discourse. *Issues in Applied Linguisitcs*, 2, 149-181.
- Jacoby, Sally & McNamara, Tim 1999. Locating competence. *English for Specific Purposes*, 18, 213-241.
- Jaeger, Richard 1990. *Statistics A Spectator Sport*. Thousand Oaks, California: Sage Publications, Inc.
- Jaffe, Alexandra M 1999. Locating power: Corsican translators and their critics. In: Blommaert, Jan (ed.) *Language Ideological Debates*. Berlin: Mouton de Gruyter.
- James, Carl 1980. *Contrastive Analysis*. Harlow Essex: Longman Group Ltd.



- Jasso-Aguilar, Rebecca 2005. Sources, methods, and triangulation in needs analysis: A critical perspective in a case study of Waikiki hotel maids. *English for Specific Purposes*, 18, 27-46.
- Jaworski, Adam & Coupland, Nikolas 2004. Sociolinguistic perspectives on metalanguage: Reflexivity, evaluation and ideology. In: Jaworski, Adam, Coupland, Nikolas & Galasinski, Dariusz (eds.) *Metalanguage*. Berlin: Gruyter de Mouton.
- Jaworski, Adam, Nikolas, Coupland & Galasinski, Dariusz (eds.) 2004. *Metalanguage*. Berlin: Gruyter de Mouton.
- Jenkins, Jennifer 2000. *The Phonology of English as an International Language: New Models, New Norms, New Goals*. Oxford: Oxford University Press.
- Jenkins, Jennifer 2006a. Points of view and blind spots: ELF and SLA. *International Journal of Applied Linguistics*, 16, 137-162.
- Jenkins, Jennifer 2006b. The spread of EIL: a testing time for testers. *ELT Journal*, 60, 42-50.
- Jenkins, Jennifer 2007. *English as a Lingua Franca: Attitude and Identity*. Oxford: Oxford University Press.
- Jenkins, Jennifer 2009. English as a lingua franca: interpretations and attitudes. *World Englishes*, 28, 200-207.
- Jenkins, Jennifer, Cogo, Alessia & Dewey, Martin 2011. Review of developments in research into English as a lingua franca. *Language Testing*, 44, 281-315.
- Jensen, Christian, Staehr, Lars Stenius & Thøgersen, Jacob 2009. *Underviseres holdninger til engelsk som undervisningssprog - en spørgeskemaundersøgelse på Københavns Universitet [Teacher Attitudes towards English as the Language of Instruction - A Survey at the University]*. University of Copenhagen: Centre for Internationalisation and Parallel Language Use.
- Jones, Jeremy F 1999. From silence to talk: Cross-cultural ideas on students' participation in academic group discussion. *English for Specific Purposes*, 18, 243-259.
- Jones, Robert L 1985. Second language performance testing: an overview. In: Hauptman, Philip C, Raymond, Leblanc & B, Wesche Marjorie (eds.) *Second language Performance Testing*. Ottawa: University of Ottawa Press.
- Juurakko, Taina (ed.) 2001. *Kielitaidon arvioinnin yhdenmukaistaminen ammattikorkeakouluissa (AMKKIA)*. Hämeenlinna: Hämeen ammattikorkeakoulu.
- Kachru, Braj B 1985. Standards, codification and sociolinguistic realism: The English language in the outer circle. In: Quirk, Randolph & Widdowson, Henry G (eds.) *English in the World: Teaching and Learning the Language and Literatures*. Cambridge: Cambridge University Press and the British Council.
- Kachru, Braj B 1986. *The Alchemy of English: The Spread, Functions, and Models of Non-native Englishes*. Oxford, England: Pergamon Press.
- Kachru, Braj B 1991. Liberation linguistics and the Quick concern. *English Today*, 7, 3-13.
- Kachru, Braj B (ed.) 1992. *The Other Tongue: English across Cultures*. Urbana: University of Illinois Press.
- Kachru, Braj B 1997. Opening borders with world Englishes: theory in the classroom. In: Cornwell, Steve, Rule, Peggy & Sugino, Toshiko (eds.) *On JALT96: Crossing Borders. Proceedings of the 23rd Annual JALT International Conference on Language Teaching and Learning*. Hiroshima, Japan: JALT Tokyo.
- Kalocsai, Karolina 2009. Erasmus exchange students: a behind-the-scenes view to an ELF community of practice. *Journal of Applied Language Studies*, 3, 25-49.
- Kane, Michael 1992. The assessment of professional competence. *Evaluation and the Health Professions*, 15, 163-182.

- Kankaanranta, Anne 2006. "Hej Seppo, Could you pls comment on this!" – Internet email. *Business Communication Quarterly*, 69, 216-225.
- Kankaanranta, Anne 2012. New winds in language education: current needs in business and engineering speak in favor of more content knowledge. *International Business Communication CLIL Seminar*. Vaasa, Finland: Unpublished paper.
- Kankaanranta, Anne & Louhiala-Salminen, Leena 2010. "English?-Oh, it's just work!": A study of BELF users' perceptions. *English for Specific Purposes*, 29, 204-209.
- Kankaanranta, Anne & Planken, Brigitte 2010. BELF competence as business knowledge of internationally operating business professionals. *Journal of Business*, 47, 308-407.
- Karjalainen, Sinikka & Lehtonen, Tuula 2005. *Että osaa ja uskaltaa kommunikoida – akateemisissa ammateissa tarvittava kielitaito työntekijöiden ja työnantajien kuvaamana*. Helsinki: Helsingin yliopiston hallinnon julkaisuja 13.
- Karlsson, Leena 1995. *Ammattien kielet ja kielten ammatit: Selvityksiä ja ehdotuksia korkeakoulujen kieltenopetuksen järjestämisestä*. Helsinki: Opetusministeriön koulutus ja tiedepolitiikan linjan julkaisusarja 33.
- Kaur, Jagdish 2009. Pre-empting Problems of Understanding in English as a Lingua Franca. In: Mauranen, Anna & Ranta, Elina (eds.) *English as a Lingua Franca: Studies and Findings*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Kelch, Ken 1985. Modified input as an aid to comprehension. *Studies in Second Language Acquisition*, 7, 81-89.
- Kim, Hyun-Ju 2006. World Englishes in language teting: a call for research. *English Today*, 88, 33-39.
- Kirkpatrick, Andy 2007. *World Englishes: Implications for International Communication and English Language Teaching*. Cambridge: Cambridge University Press.
- Kirkpatrick, Andy 2011. Internationalization or Englishization: Medium of Instruction in Today's Universities. In: Citizenship, Centre for Governance and Education (ed.). The Hong Kong Institute of Education: Editor of the CGC Working Paper Series: Dr Betty Yung.
- Kivistö, Jussi & Tirronen, Jarkko 2012. New elitism in universal higher education: The building process, policy and idea of Aalto University. In: Ahola, Sakari & Hoffman, David (eds.) *Higher Education Research in Finland*. Jyväskylä, Finland: University of Jyväskylä, Finnish Institute for Educational Research.
- Klaassen, Renate 2001. Preparing lectures for English-medium instruction. In: Wilkinson, Robert & Zegers, Vera (eds.) *Realizing Content and Language Integration in Higher Education*. Maastricht, Netherlands: Maastricht University Language Centre.
- Klaassen, Renate G. & Bos, Madeleine 2010. English language screening for scientific staff at Delft University of Technology. *HERMES – Journal of Language and Communication Studies*, 45, 62-75.
- Kling, Joyce & Staehr, Lars Stenius 2011. Assessment and assistance: Developing university lecturers' skills through certification feedback. In: Cancino, Rita, Dam, Lotte & Jaeger, Kirsten (eds.) *Policies, Principles, Practices: New Directions in Foreign Language Education in the Era of Educational Globalization*. Newcastle upon Tyne: Cambridge Scholars.
- Knapp, Karlfried & Meierkord, Christiane 2002. Approaching lingua franca communication. In: Knapp, Karlfried & Meierkord, Christiane (eds.) *Lingua Franca Communication*. Frankfurt am Main: Peter Lang.
- Kramsch, Claire 1993. *Context and Culture in Language Teaching*. Oxford: Oxford University Press.
- Kroskrity, Paul V. 1998. Arizona Tewa Kiva speech as a manifestation of dominant language ideology. In: Scheffelin, Bambi B., Woolard, Kathryn A. & Kroskrity, Paul V. (eds.) *Language Ideologies: Theory and Practice*. Oxford: Oxford University Press.

- Kroskrity, Paul V. (ed.) 2000. *Regimes of Language*. Santa Fe, NM: School of American Research Press.
- Kroskrity, Paul V. 2004. Language ideologies. In: Duranti, Alessandro (ed.) *A Companion to Linguistic Anthropology*. Oxford: Blackwell.
- Kvale, Steinar 1996. *InterViews – An Introduction to Qualitative Research Interviewing*. London: Sage.
- Labov, William 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- Lado, Robert 1961. *Language Testing: The Construction and Use of Foreign Language Tests*. London: Longman, Green, and Company Limited.
- Lafromboise, Teresa , Coleman, Hardin L & Gerton, Jennifer 1993. Psychological impact of biculturalism: Evidence and theory. *Psychological Bulletin*, 14, 395-419.
- Lappalainen, Pia (ed.) *It's Just People with People: Views of Corporate Social Responsibility*. Espoo, Finland: Aalto University Press.
- Laskey, Marcia L & Hetzel, Carole J 2010. Self-regulated learning, metacognition and soft skills: the 21st century learner [Online]. Online submission, 2010 ERIC. Available at: [eric.ed.gov:80/PDFS/ED511589.pdf](http://eric.ed.gov:80/PDFS/ED511589.pdf). [Accessed July 31, 2012 2012].
- Lave, Jean & Wenger, Etienne 1991. *Situated learning: Legitimate Peripheral Participation*. New York: Cambridge University Press.
- Lawton, William & Katsomitros, Alex. 2012. International branch campuses expanding, geopolitical landscape changing. *University World News*, January 22, 2012.
- Lazaraton, Ann 2005. Quantitative research methods. In: Hinkel, Eli (ed.) *Handbook of Research in Second Language Teaching and Learning*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Lee, Joseph J 2009. Size matters: an exploratory comparison of small- and large-class university lecture introductions. *English for Specific Purposes*, 28, 42-57.
- Lennon, Paul 2000. The lexical element in spoken second language fluency. In: Riggensbach, Heidi (ed.) *Perspectives on Fluency*. Ann Arbor, MI: University of Michigan Press.
- Leung, Constant & Lewkowicz, Jo 2006. Expanding Horizons and Unresolved Conundrums: Language Testing and Assessment. *TESOL Quarterly*, 40, 211-234.
- Levett-Jones, Tracy 2005. Self-directed learning: Implications and limitations for undergraduate nursing education. *Nurse Education Today*, 25, 363-368.
- Levis, John M 2006. Pronunciation and the assessment of spoken language. In: Hughes, Rebecca (ed.) *Spoken English, TESOL, and Applied Linguistics: Challenges for Theory and Practice*. Houndsmills, Basingstoke: Palgrave MacMillan.
- Lincoln, Yvonna S & Guba, Egon G 1985. *Naturalistic Enquiry*. Newbury Park, CA: Sage Publications.
- Lippi-Green, Rosina 1997. *English with an Accent: Language, Ideology, and Discrimination in the United States*. London: Routledge.
- Little, David 2006. The Common European Framework of Reference for Languages: content, purpose, origin, reception, and impact. *Language Teaching*, 39, 167-190.
- Long, Michael 2005. *Second Language Needs Analysis*. Cambridge, England: Cambridge University Press.
- Lopes, Marilyn. 1992. *Native Speaker Performance on the IELTS Reading Test*. University of Melbourne.
- Louhiala-Salminen, Leena & Kankaanranta, Anne 2011. Professional communication in a global business context: The notion of global communicative competence. *IEEE Transactions on Professional Communication*, 54, 244-262.

- Lumley, Tom & Brown, Annie 1998. Authenticity of discourse in a specific purpose test. In: Li, Edward & James, Gregory (eds.) *Testing and Evaluation in Second Language Education*. Hong Kong: Language Centre, Hong Kong University of Science and Technology.
- Luoma, Sari 2004. *Assessing Speaking*. Cambridge, England: Cambridge University Press.
- Lyle, John 2003. Stimulated recall: a report on its use in naturalistic research. *British Educational Research Journal*, 29, 861-878.
- Lynch, Michael 1985. *Art and Artefact in Laboratory Life: A Study of Shop Work and Shop Talk in a Research Laboratory*. London: Routledge and Kegan Paul.
- Mackey, Alison & Gass, Susan M 2005. *Second Language Research: Methodology and Design*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Majanen, Silke. 2008. *English as a Lingua Franca: Teachers' Discourse on Accent and Identity*. M.A. thesis, University of Helsinki.
- Marlaire, Courtney L & Maynard, Douglas W 1993. Social problems and the organization of talk and interaction. In: Holstein, James A & Miller, Gale (eds.) *Reconsidering Social Constructionism: Debates in Social Problems Theory*. New York: de Gruyter.
- Marx, Nicole 2002. Never quite a 'native speaker': Accent and identity in the L2 - and the L1. *The Canadian Modern Language Review*, 59, 264-281.
- Maslen, Geoff. 2009. Global: Huge expansion in overseas campuses. *University World News Global Edition*, November 22, 2009.
- Mauranen, Anna 2003. The corpus of English as a lingua franca in academic settings. *TESOL Quarterly*, 37, 513-527.
- Mauranen, Anna 2005a. English as a lingua franca – an unknown language? In: Cortese, Giuseppina & Duszak, Anna (eds.) *Identity, Community, Discourse: English in Intercultural Settings*. Frankfurt: Peter Lang.
- Mauranen, Anna 2005b. Speaking academics. In: Bäcklund, Ingegerd, Börestam, Ulla, Marttala, Ulla Melander & Näslund, Harry (eds.) *Text i arbete/text at work*. Uppsala: ASLA.
- Mauranen, Anna 2006a. Signaling and preventing misunderstanding in English as a lingua franca communication. *International Journal of the Sociology of Language*, 177, 123-150.
- Mauranen, Anna 2006b. Spoken discourse, academics and global English: A corpus perspective. In: Hughes, Rebecca (ed.) *Spoken English, Applied Linguistics and TESOL: Challenges for Theory and Practice*. Houndmills, Basingstoke: Palgrave Macmillan.
- Mauranen, Anna 2007a. Hybrid voices: English as the lingua franca of academics. In: Fløttum, Kjersti (ed.) *Language and Discipline Perspectives on Academic Discourse*. Newcastle-upon-Tyne: Cambridge Scholars.
- Mauranen, Anna 2007b. *Lecture series. Spoken language and ELF*. Department of Modern Languages, University of Helsinki.
- Mauranen, Anna 2009a. Chunks in ELF: Expressions for managing interaction. *Journal of Intercultural Pragmatics*, 6, 217-233.
- Mauranen, Anna 2009b. Spoken rhetoric: How do natives and non-natives fare? In: Suomela-Salmi, Eija & Dervin, Fred (eds.) *Cross-linguistic and Cross-cultural Perspective on Academic Discourse*. Amsterdam: John Benjamins.
- Mauranen, Anna 2010. Discourse reflexivity-A discourse universal? The case of ELF. *Nordic Journal of English Studies*, 9, 13-40.
- Mauranen, Anna 2011a. English as the lingua franca of the academic world. In: Bpaltridge, Brian, Belcher, Diane & Johns, Ann (eds.) *New Directions in English for Specific Purposes Research*. Ann Arbor: University of Michigan Press.

- Mauranen, Anna 2011b. Learners and users - Who do we want corpus data from? In: Meunier, Fanny, Cock, Sylvie De, Gilquin, Gaetanelle & Paquot, Magali (eds.) *A Taste for Corpora. In honour of Sylviane Granger*. Amsterdam: John Benjamins Publishing Company.
- Mauranen, Anna 2012. *Exploring ELF: Academic English Shaped by Non-native Speakers*. Cambridge: Cambridge University Press.
- Maxwell, Joseph A 1992. Understanding the validity in qualitative research. *Harvard Educational Review*, 62, 279-300.
- McClelland, David C 1973. Testing for competence rather than for "intelligence". *American Psychologist*, 28, 1-14.
- McNamara, Tim 1989. ESP testing: general and particular. In: Candlin, Christopher & McNamara, Tim (eds.) *Language, Learning, and Community*. Sydney, NSW: National Centre for English Language Teaching and Research, Macquarie University.
- McNamara, Tim 1996. *Measuring Second Language Performance*. Harlow, Essex, UK: Addison Wesley Longman Ltd.
- McNamara, Tim 2001. Language assessment as social practice: Challenges for research. *Language Testing*, 18, 333-349.
- McNamara, Tim 2005. 21st century Shibboleth: language tests, identity and intergroup conflict. *Language Policy*, 4, 1-20.
- McNamara, Tim & Roever, Carsten 2006. *Language Testing: The Social Dimension*. Oxford, UK: Blackwell Publishing.
- Medgyes, Peter 1994. *The Non-native Teacher*. London: Macmillan.
- Meierkord, Christiane 2000. Interpreting successful lingua franca interaction: An analysis of non-native-non-native small talk conversations in English. In: Pittner, Karin & Anita, Fetzer (eds.) *Neuere Entwicklungen in Der Gesprächsforschung. Sonderausgabe Von Linguistic*. Available online: [http://www.linguistik-online.de/1\\_00/MEIERKOR.HTM](http://www.linguistik-online.de/1_00/MEIERKOR.HTM) [Accessed 12 August 2011].
- Meisel, Perry & Saussy, Haun (eds.) 1959. *Course in General Linguistics: Ferdinand de Saussure*. New York: Columbia University Press.
- Messick, Samuel 1989. Validity. In: Linn, Robert L (ed.) *Educational Measurement*. Third edition. New York: Macmillan.
- Messick, Samuel 1994. The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23, 13-23.
- Miller, Lindsay 2009. Engineering lectures in a second language: what factors facilitate students' listening comprehension? *The Asian EFL Journal Quarterly*, 11, 8-30.
- Milroy, James & Milroy, Leslie 1999. *Authority in Language*, London, Routledge.
- Milroy, Leslie 1999. Standard English and language ideology in Britain and the United States. In: Bex, Toni & Watts, Richard J. (eds.) *Standard English: The Widdening Debate*. London: Routledge.
- Morell, Teresa 2004. Interactive lecture discourse for university EFL students. *English for Specific Purposes*, 23, 325-338.
- Morrow, Keith 1979. Communicative language testing: revolution or evolution? In: Brumfit, Christopher & Johnson, Keith (eds.) *The Communicative Approach to Language Teaching*. Oxford: Oxford University Press.
- Morrow, Susan L 2005. Quality and trustworthiness in qualitative research in counseling psychology. *Journal of Counseling Psychology*, 52, 250-260.
- Morse, J.M. 2003. Principles of mixed methods and multimethod research design. In: Tashakkori, A & Teddlie, C. (eds.) *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks, CA: Sage.

- Mortensen, Janus 2008. 'Circus English' Investigating English as an academic lingua franca in BA study group meetings at Roskilde University. *Higher Education in the Global Village: Cultural and Linguistic Practices in the International University*, 85-95.
- Mortimer, Colin 1985. *Elements of Pronunciation: Intensive Practice for Intermediate and More Advanced Students*. Cambridge: Cambridge University Press.
- Mufwene, Salikoko S. 2001. *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.
- Mulligan, Denise & Kirkpatrick, Andy 2000. How much do they understand? Lectures, students, and comprehension. *Higher Education Research and Development*, 19, 311-335.
- Nash, Rose 1969. Intonational interference in the speech of Puerto Rican bilinguals. *Journal of English as a Second Language*, 4, 1-42.
- Nattinger, James R & Decarrico, Jeanette S 1992. *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.
- Nederström, Mikae & Niitamo, Petteri 2010. Construction and validation of a work personality inventory. *Aalto Publication Series*. Espoo, Finland: Aalto University.
- Newman, Isadore & Benz, Carolyn R 1998. *Qualitative-quantitative Research Methodology: Exploring the Interactive Continuum*. Chicago: Southern Illinois University.
- Niitamo, Petteri 1999. Evaluation for Cross-Cultural Work: Assessment Centre Methods in Finland. *Cross-Cultural Psychology Bulletin*.
- Niitamo, Petteri 2011. Intercultural work competence. . In: Lappalainen, Pia (ed.) *It's Just People with People: Views of Corporate Social Responsibility*. Espoo, Finland: Aalto University Press.
- Nikko, Tuija 2005. *CEF as a Tool for Evaluating Foreign Language Communication Competence at Finnish Universities*. Helsinki: University of Helsinki Language Centre.
- Nousiainen, Anu. 2009. Mai neim is Mäti. *Helsinginsanommat Kuukausiliite*.
- O'malley, J. Michael & Pierce, Lorraine Valdez 1996. *Authentic Assessment for English Language Learners: Practical Approaches for Teachers*. USA: Addison-Wesley Publishing Company.
- O'sullivan, Barry. 2000. *Towards a Model of Performance in Oral Language Testing*. Doctor of Philosophy, University of Reading, England.
- Ochs, Elinor 1988. *Culture and Language Development: Language Acquisition and Language Socialization in a Samoan village*. Cambridge, UK: Cambridge University Press.
- Oller, John W Jr 1979. *Language Tests at School: A Pragmatic Approach*. London: Longman.
- Opm 2007. *Hallituksen iltakoulu 27.11.2007: Yliopistojen taloudellisen ja hallinnollisen aseman uudistaminen ja innovaatioyliopiston perustaminen [Government night school 27.11.2007: Reforming the financial and administrative position of universities and establishing an innovation university]*. Helsinki: Finnish Ministry of Education.
- Opm 2009a. *Strategy for Internationalization Activities in Higher Education Institutions in Finland 2009-2015*. Helsinki: Finnish Ministry of Education.
- Opm 2009b. *Yliopistot ja ammattikorkeakoulut korkeakoulupolitiikan toteuttajina [Universities and polytechnics as implementers of higher education policy]*. Helsinki: Opetusministeriön julkaisu 2009,49. The New Universities Act, 558/2009.
- Overstreet, Maryann & Yule, George 1999. Fostering L2 pragmatic awareness. *Applied Language Learning* 10, 1-14.

- Paikeday, Thomas M 1985. *The Native Speaker is Dead!* Toronto and New York: Paikeday.
- Pavlenko, Aneta & Blackledge, Adrian (eds.) 2004. *Negotiation of Identities in Multilingual Contexts*. Clevedon: Multilingual Matters.
- Pawley, Andrew & Syder, Frances H 1983. Two puzzles for linguistic theory: natively-like selection and natively-like fluency. In: Richards, Jack C & Schmidt, Richard W (eds.) *Language and Communication*. London: Longman.
- Pennycook, Alastair 1994. *The Cultural Politics of English as an International Language*. Harlow, Essex, UK: Longman Group Limited.
- Phillipson, Robert 1992. *Linguistic Imperialism*. Oxford: Oxford University Press.
- Pihko, Marja-Kaisa 2004. When does 'nice' become 'awful'? Target language attitude development during school language learning. In: Kaikkonen, Pauli, Mäkinen, Kaarina & Korhonen, Viljo (eds.) *Future Perspectives in Foreign Language Education*. Oulu: University of Oulu.
- Pilkinton-Pihko, Diane 2010. English as a lingua franca lecturers' self-perceptions of their language use. *Helsinki English Studies*, 6, 58-74.
- Pilkinton-Pihko, Diane 2011. *Lecturer Attitudes Towards and Perceptions of Teaching in English as a Lingua Franca*. Crossover. Espoo: Aalto University.
- Pitzel, Marie-Luise (ed.) 2009. *"We should not wake up any dogs": Idiom and Metaphor in ELF*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Pözl, Ulrike & Seidlhofer, Barbara 2006. In and on their own terms: The 'habitat factor' in English as a lingua franca interactions. *International Journal of the Sociology of Language*, 177, 151-176.
- Prator, Clifford 1968. The British heresy in TESL. In: Fishman, Joshua, Ferguson, Charles & Das Gupta, Jyotirinda (eds.) *Language Problems in Developing Nations*. New York: John Wiley.
- Psacharopoulos, George 1973. *Returns to Education: An International Comparison*. Amsterdam: Elsevier.
- Quirk, Randolph 1985. The English language in a global context. In: Quirk, Randolph & Widdowson, Henry G (eds.) *English in the World: Teaching and Learning the Language and Literatures*. Cambridge: Cambridge University Press.
- Quirk, Randolph 1990. Language varieties and standard language. *English Today*, 21, 3-10.
- Rajadurai, Joanne 2006. Pronunciation issues in non-native contexts: A Malaysian case study. *Malaysian Journal of ELT Research*, 2, 42-59.
- Rampton, Ben 2007. Neo-Hymesian linguistic ethnography in the United Kingdom. *Journal of Sociolinguistics*, 11, 584-607.
- Ranta, Elina. 2004. *International English - A Future Possibility in the Finnish EFL Classroom?* Master of Arts thesis, University of Tampere.
- Ranta, Elina 2006. The 'attractive' progressive - why use the -ing form in English as a lingua franca? *Nordic Journal of English Studies*, 5, 95-116.
- Ranta, Elina 2009. Syntactic Features in Spoken ELF - Learner Language or Spoken Grammar? In: Mauranen, Anna & Ranta, Elina (eds.) *English as a Lingua Franca: Studies and Findings*. Newcastle: Cambridge Scholars Publishing.
- Ranta, Elina 2010. English in the real world vs English at school - Finnish English teachers' and students' views. *International Journal of Applied Linguistics*, 20, 156-177.
- Ranta, Elina. 2013. *Universals in a Universal Language? - Exploring Verb-Syntactic Features in English as a Lingua Franca*. Doctor of philosophy, University of Tampere.

- Richterich, Rene & Chancerel, Jean-Louis 1977. *Identifying the needs of adults learning a foreign language*. Oxford, England: Pergamon Press and Council of Europe.
- Rogerson-Revell, Pamela 2010. "Can you spell that for us non-native speakers?" Accommodation strategies in international business meetings. *Journal of Business Communication*, 47, 432-454.
- Rontu, Heidi 2010. Kieli teknologiatyössä – diplomi-insinöörejä rekrytoivien yritysten kuvauksia henkilöstön kieli- ja viestintätaitotarpeista. *Vaasa: VAKKI:n julkaisut*, N:o 37 [Online] [http://www.vakki.net/publications/2010/VAKKI2010\\_Rontu.pdf](http://www.vakki.net/publications/2010/VAKKI2010_Rontu.pdf) [Accessed: 30 November 2012].
- Rumsey, Alan 1990. Wording, Meaning, and Linguistic Ideology. *American Anthropologist*, 92, 346-361.
- Saarinen, Taina 2008. Whose quality? Social actors in the interface of transnational and national higher education policy. *Discourse: Studies in Cultural Politics of Education* 29, 179-193.
- Saarinen, Taina 2012. Internationalization and the invisible language? Historical phases and current policies in Finnish higher education. In: Ahola, Sakari & Hoffman, David (eds.) *Higher Education Research in Finland: Emerging Structures and Contemporary Issues*. Jyväskylä, Finland: University of Jyväskylä: Finnish Institute for Educational Research.
- Sajavaara, Kari 1986. Transfer and second language speech processing. In: Kellerman, Eric & Sharwood Smith, Michael (eds.) *Crosslinguistic Influence in Second Language Acquisition*. New York: Pergamon Press.
- Saldana, Johnny 2009. *The Coding Manual for Qualitative Researchers*. London: Sage Publications.
- Sale, Joanna E. M, Lohfeld, Lynne H & Brazil, Kevin 2002. Revisiting the quantitative-qualitative debate: Implications for mixed-methods research. *Quality and Quantity*, 36, 43-53.
- Sallinen, Noora. 2009. "Poison or Spice in the Language?" - A Study of Attitudes of Finnish University Students of English Philology towards Different Accents of English. M.A. thesis, University of Tampere.
- Savignon, Sandra 1972. *Communicative Competence: An Experiment in Foreign Language Teaching*. Philadelphia PA: The Center for Curriculum Development.
- Savignon, Sandra 1983. *Communicative Competence: Theory and Classroom Practice*. Reading, MA: Addison-Wesley.
- Schegloff, Emanuel A 1982. Discourse as an interactional achievement: Some uses of "uh huh" and other things that come between sentences. In: Tannen, Deborah (ed.) *Georgetown University Roundtable on Languages and Linguistics*. Washington, DC: Georgetown University Press.
- Schegloff, Emanuel A 1988. Discourse as an interactional achievement II: An exercise in conversation analysis. In: Tannen, Deborah (ed.) *Linguistics in Context: Connecting Observation and Understanding*. Norwood NJ: Ablex.
- Schegloff, Emanuel A 1995. Discourse as an interactional achievement III: The omnirelevance of action. *Research on Language and Social Interaction*, 28, 185-211.
- Schieffelin, Bambi B 1990. *The Give and Take of Everyday Life: Language Socialization of Kaluli Children*. Cambridge, UK: Cambridge University Press.
- Schieffelin, Bambi B & Doucet, Rachelle C 1998. The "Real" Haitian Creole: Ideology, metalinguistics and orthographic choice. In: Schieffelin, Bambi B, Woolard, Kathryn A & Kroskrity, Paul V (eds.) *Language Ideologies: Practice and Theory*. Oxford and New York: Oxford University Press.
- Schieffelin, Bambi B & Ochs, Elinor (eds.) 1986. *Language Socialization across Cultures*. Cambridge, UK: Cambridge University Press.



- Schieffelin, Bambi B, Woolard, Kathryn A & Kroskrity, Paul V (eds.) 1998. *Language Ideologies: Practice and Theory*. Oxford: Oxford University Press.
- Schön, Donald 1983. *The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books.
- Schultz, Theodore W 1961. Investment in Human Capital. *American Economic Review*, 51, 1-17.
- Sergeant, Phillip 2009. *The Idea of English in Japan*. Bristol: Multilingual Matters.
- Seidlhofer, Barbara 2000. Mind the gap: English as a mother tongue versus English as a lingua franca, *Views*, 9, 51-68.
- Seidlhofer, Barbara 2001. Closing a conceptual gap: the case for a description of English as a lingua franca. *International Journal of Applied Linguistics*, 11, 133-158.
- Seidlhofer, Barbara. 2003. *A Concept of international English and related issues: from 'real' English to 'realistic English'?* [Online]. Strasbourg: Council of Europe. Available: <http://www.coe.int/t/dg4/linguistic/Source/SeidlhoferEN.pdf> [Accessed 30 November 2006].
- Seidlhofer, Barbara 2004. Research perspectives on teaching English as a lingua franca. *Annual Review of Applied Linguistics*, 24, 209-239.
- Seidlhofer, Barbara 2011. *Understanding English as a Lingua Franca*. Oxford: Oxford University Press.
- Seidlhofer, Barbara & Widdowson, Henry (eds.) 2003. House work and student work: A study in cross-cultural understanding, 8 (2/3), 115-127. Available online: [http://zif.spz.tu-darmstadt.de/jg-08-2-3/docs/Seidlhofer\\_Widdowson.pdf](http://zif.spz.tu-darmstadt.de/jg-08-2-3/docs/Seidlhofer_Widdowson.pdf) [Accessed 19 April 2013].
- Selinker, Larry 1972. Interlanguage. *International Review of Applied Linguistics*, 10, 209-231.
- Selinker, Larry 1979. On the use of informants in discourse analysis and language for specific purposes. *International Review of Applied Linguistics*, 17, 189--215.
- Selinker, Larry 1992. *Rediscovering Interlanguage*. London: Longman.
- Shannon, Sheila M 1999. The debate on bilingual education in the U.S.: Language ideology as reflected in the practice of bilingual teachers. In: Blommaert, Jan (ed.) *Language Ideological Debates*. Berlin: Mouton de Gruyter.
- Shohamy, Elana 2001. Democratic testing as an alternative. *Language Testing*, 18, 373-391.
- Silverstein, Michael 1979. Language structure and linguistic ideology. In: Clyne, Paul R., Hanks, William F. & Hofbauer, Carol L. (eds.) *The Elements: A Parasession on Linguistic Units and Levels*. Chicago: Chicago Linguistic Society.
- Simon, Roger 1982. *Gramsci's Political Thought: An Introduction*. London: Lawrence and Wishart.
- Sinclair, John & Mauranen, Anna 2006. *Linear Unit Grammar*. Amsterdam: John Benjamins.
- Singh, Rajendar (ed.) 1998. *The Native Speaker: Multilingual Perspectives*. New Delhi: Sage.
- Sinkkonen, Marjukka 1998. Kielitaidon tarvetutkimukset työelämässä. In: Takala, Sauli & Saajavaara, Kari (eds.) *Kielikoulutus Suomessa*. Jyväskylä: Jyväskylän yliopisto soveltavan kielentutkimuksen keskus.
- Smit, Ute. 2008. *Classroom Interaction in English as a Lingua Franca: A Discourse Pragmatic Ethnography in International Higher Education*. University of Vienna.
- Smit, Ute 2010. CLIL in an English as a lingua franca (ELF) classroom: On explaining terms and expressions interactively. In: Dalton-Puffer, Christiane, Nikula, Tarja &

- Smit, Ute (eds.) *Language Use and Language Learning in CLIL Classrooms*. Amsterdam: John Benjamins.
- Smith, Jonathan A. & Osborn, Mike 2008. Interpretive phenomenological analysis. In: Smith, Jonathan A (ed.) *Qualitative Psychology: A Practical Guide to Research Methods*. London: Sage Publishing
- Sobkowiak, Włodzimierz 2005. Why not LFC? In: Kdziubalska-Kolaczyk, Katarzyna & Przedlacka, Joanna (eds.) *English Pronunciation Models: A Changing Scene*. Frankfurt: Peter Lang.
- Söderlundh, Hedda. 2010. *Internationella universitet - lokala språkval Om bruket av talad svenska i engelskspråkiga kursmiljöer / International Universities - Local Language Choices On Spoken Swedish in English-medium Course Environments*. Doctor of philosophy, Uppsala Universitet.
- Spensor, Lyle M & Spencer, Signe M. 1993. *Competence at Work: Models for Superior Performance*. New York: John Wiley and sons.
- Spolsky, Bernard 1968. Language testing: the problem of validation. *TESOL Quarterly*, 2, 88-94.
- Spradley, James 1979. *The Ethnographic Interview*. New York: Holt, Rinehart and Winston.
- Strodt-Lopez, Barbara 1991. Tying it all in: Asides in University Lectures. *Applied Linguistics*, 12, 117-140.
- Suviniitty, Jaana 2010. Lecturers' questions and student perception of lecture comprehension. *Helsinki English Studies*, 6, 44-57.
- Suviniitty, Jaana. 2012. *Lectures in English as a Lingua Franca: Interactive Features*. Doctor of Philosophy, University of Helsinki.
- Swan, Michael, Smith, Bernard & Ur, Penny 2001. *Learner English: A Teacher's Guide to Interference and Other Problems*. Cambridge: Cambridge University Press.
- Sweeney, Emma & Hua, Zhu 2010. Accommodating toward your audience. Do native speakers of English know how to accommodate their communication strategies toward nonnative speakers of English? *Journal of Business Communication*, 47, 477-504.
- Szpyra-Kozłowska, Jolanta 2003. The lingua franca core and the Polish learner. *Neofilologia*, 5, 193-210.
- Takala, Sauli 2001. Ammattikorkeakoulujen opiskelijoiden kielitaidon arviointi-hankkeen koespesifikaatiot. In: Juurakko, Taina (ed.) *Kielitaidon arvioinnin yhdenmukaistaminen ammattikorkeakouluissa*. Häme: Hämeen ammattikorkeakoulun julkaisu A:16 (29-123).
- Tashakkori, Abbas & Teddlie, Charles (eds.) 2003. *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks, CA: Sage Publications.
- Taylor, Lynda 2006. The changing landscape of English: implications for language assessment. *ELT Journal*, 60, 51-60.
- Toepfer, Tom & Virkkunen-Fullenwider. 2005. *Introduction of CEF as a tool for Evaluation and Training of Teachers in its Use* [Online]. Helsinki: Helsinki University Language Centre. Available: <http://www.helsinki.fi/kksc/cef2005/documents/ToepferVirkkunen.pdf> [Accessed 29 June 2010].
- Treaty, Lisbon 2000. *Lisbon European Council: Presidency Conclusions*. [http://www.europarl.europa.eu/summits/lis\\_1\\_en.htm](http://www.europarl.europa.eu/summits/lis_1_en.htm) [Accessed 17 April 2012].
- Treaty, Maastricht 1992. *Treaty on the European Union*. Maastricht, The Netherlands: <http://eur-lex.europa.eu/en/treaties/dat/11992M/htm/11992M.html> [Accessed 17 April 2012].
- Trudgill, Peter 1974. *An Introduction to Sociolinguistics*. Harmondsworth, Middlesex: Penguin Books.

- Tyler, Andrea 1995. The coconstruction of miscommunication. *Studies in Second Language Acquisition*, 17, 129-152.
- University, Aalto 2010. *Aalto University Language Guidelines*. Helsinki, Finland: Aalto University.
- University, Aalto 2011. *Implementation Plan for Language Guidelines: 2012-2016*. Helsinki, Finland: Aalto University.
- Urciuoli, Bonnie 1991. The political topography of Spanish and English: The view from a New York Puerto Rican neighborhood. *American Ethnologist*, 295-310.
- Urciuoli, Bonnie 1996. *Exposing Prejudice: Puerto Rican Experiences of Language, Race, and Class*. Boulder Col: Westview.
- Van Hest, Erna & Oud De Glas, Maria 1991. *A Survey of Techniques used in the Diagnosis and Analysis of Foreign Language Needs in Trade and Industry*. Nijmegen, Institute of Applied Linguistics and Research Methodology.
- Van Wijngaarden, Sandra J 2001. Intelligibility of native and non-native Dutch speech. *Speech Communication*, 35.
- Van Wijngaarden, Sandra J, Steeneken, Herman J M & Houtgast, Tammo 2002. Quantifying the intelligibility of speech in noise for non-native listeners. *Journal of Acoustical Society of America*, 111, 1906-1916.
- Vandermeeren, Sonja 2005. Foreign language need of business firms. In: Long, Michael (ed.) *Second Language Needs Analysis*. Cambridge, England: Cambridge University Press.
- Vestergaard, Torben 1996. Writing English grammars for Danes: Bridging the gap between prescriptivism and descriptivism. *Journal of Linguistics*, 17, 113-121.
- Vinke, Andriana A. 1995. *English as the Medium of Instruction in Dutch Engineering Education*. Doctor of Philosophy Doctoral dissertation, Technische Universiteit Delft.
- Vinke, Andriana A., Snippe, Joke & Jochems, Wim 1998. English-medium content courses in non-English higher education: a study of lecturer experiences and teaching behaviors. *Teaching in Higher Education*, 3, 383-394.
- Vygotsky, Lev S 1978. *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.
- Wächter, Bernd & Maiworm, Friedhelm 2008. *English-Taught Programmes in European Higher Education: The Picture in 2007*. Bonn: Lemmens.
- Warren, Christopher 2006. *Features of Naturalness in Conversation*. Amsterdam: John Benjamins.
- Watts, Richard J. 1999. *The Ideology of Dialect in Switzerland: Language Ideological Debates*. Berlin: Mouton de Gruyter.
- Widdowson, Henry 1994. The Ownership of English. *TESOL Quarterly*, 28, 377-389.
- Widdowson, Henry 1998a. Context, Community, and Authentic Language. *TESOL Quarterly*, 32, 705-716.
- Widdowson, Henry 1998b. Skills, abilities, and contexts of reality. *Annual Review of Applied Linguistics*, 18, 323-333.
- Wigglesworth, Gillian 1997. An investigation of planning time and proficiency level on oral test discourse. *Language Testing*, 14, 85-106.
- Witteman, Marijt J, Weber, Andrea & Mcqueen, James M 2010. Rapid and long-lasting adaptation to foreign-accented speech. *Journal of Acoustical Society of America*, 128, 2486.
- Woolard, Kathryn A. 1998. Introduction: Language Ideology as a Field of Inquiry. In: Scheffelin, Bambi B., Woolard, Kathryn A. & Kroskrity, Paul V. (eds.) *Language Ideologies: Theory and Practice*. Oxford: Oxford University Press.
- Wray, Allison 2002. *The Transition to Language*. Oxford: Oxford University Press.

Zimmerman, Barry J 1990. Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25, 3-17.

## Appendix A: Self-analysis

Mentee's name

Native language

### Background

What is your background with the English language? Where did you learn it? And how actively have you been using it for academic and professional purposes?

### My teaching experience

When?	Where?	Subject?	What language?

- Other work experience where the working language was English? Please explain.
- Why are you taking part in this mentoring program?
- What would you like to achieve during the mentoring process?
- In relation to what you would like to achieve, what goals have you set for yourself?

### Evaluate your skills in the following areas:

*For academic and professional settings, how do you rate your own language ability in English for each of the following areas? Put an X in the appropriate box.*

	Less than adequate	adequate	good	excellent
Listening				
Speaking				
Writing				
Subject-specific terminology				

How do you rate your English language ability for *everyday communication in non-professional settings*? Put an X in the appropriate box.

	Less than adequate	adequate	good	excellent
Listening				
Speaking				
Writing				
Subject-specific terminology				

Please rank order the following items in terms of your needs, such that 1 = the skill that needs training most, and 10 = the skill that needs training least.

For example: 4 Speaking, pronunciation

7 Presentation skills

	Speaking, fluency
	Speaking, pronunciation
	Presentation skills
	Cultural awareness
	Subject-specific terminology
	Grammar
	Pedagogical skills
	Listening skills
	Writing skills
	Other, please specify:

*When you have completed this profile, email it back to your mentor [firstname.lastname@aalto.fi](mailto:firstname.lastname@aalto.fi) and be prepared to discuss these items at our first meeting.*

## **Appendix B: Guide for interview 1**

### **Interview 1: Discuss profile and self-analysis (see Appendix A)**

#### **Language background**

1. Could you begin by briefly telling me about your background as a learner of English? Tell me about how your English teacher taught English, and what was considered important. And how about nowadays, what do you feel is important when you use English for lecturing?
2. How many years did you study English at school? How about other languages?
3. Tell me about your background as an educator/researcher?
4. In what contexts have you used your English?
5. Have you taught in other institutions than Aalto?
6. Since when have you been lecturing in English?
7. Tell me about your experiences in teaching through English. What has that been like? What was it like in the beginning? What's it like now? What seems easy? Challenging?
8. Tell me about your other professional experiences where you've been using English in international contexts. What kind of situations? Easy? Challenging? How are those situations different from/similar to the classroom?
9. Any previous training in working in multicultural environments?

#### **Self-assessment**

10. What are your general feelings about your own English?
11. Self-analysis – what differences between using in non-professional vs professional situations?
12. Self-analysis – strong and weak areas? (review rankings on self-assessment form)
13. In terms of teaching in English, what do you feel that you need help with?
14. In terms of language and communication, what do you feel that you need help with? (Are there some things that you feel uncertain about?)
15. Is there something specific that you would like for me to focus on?

## **Appendix C: Guide for interview 2**

### **Interview 2: Semi-structured interview following lecture observation**

#### **General follow-up questions**

1. How are you today? Anything new?
2. How did you feel today when lecturing? Excited? Nervous? Relaxed? Please explain.
3. What's your general impression of the classroom situation?
4. How about your language use in the classroom? How does teaching in English feel?
5. What went well in today's lecture? What didn't?
6. At any point during the lecture, do you remember encountering any problems?
7. How about in some other recent lectures, can you recall any linguistic difficulties that you've encountered while lecturing? How did you overcome them?
8. Thinking about this lecture, what do you feel needs improvement?

#### **Interactional aspect**

1. How do you know the students are following the ideas and concepts?
2. What do you think is important for clarity in lectures?
3. What could be done to increase activation (with the students)?
4. During the lecture, a couple of students had questions. Could you comment on what kinds of questions? How did interacting with the student go?

#### **Image**

1. What is your image of yourself as a professional lecturer in English? How do you see yourself?
2. What do you think is important when lecturing? How do you hope your students view you?

#### **Standards/targets**

1. How would you describe your professional spoken English?
2. Generally, to lecture well in English, what do you feel you need to be able to do to?
3. What kinds of standards have you set for yourself? What is your main goal when lecturing? How about language-wise?
4. What are your views on correctness in spoken English?
5. What is your idea of correct grammar?

#### **Accent**

1. If you could have any English accent, what would you choose?
2. How would you describe the accent of a native speaker of English?
3. How would describe your accent in English?
4. How do you feel about your accent in English? Or do you like your accent in English?
5. To be a successful lecturer in English here in Finland, what kind of accent do you think would be a good target?



## **Appendix D: Guide for interview 3**

### **Interview 3: CEFR scales and guidelines for semi-structured interview**

#### **Interviewees were shown the following CEFR scales**

1. Common reference levels: global scale (Europe 2001:24)
2. Vocabulary range (Europe 2001:112)
3. Grammatical accuracy (Europe 2001:114)
4. Phonological control (Europe 2001:117)
5. Coherence and cohesion (Europe 2001:125)
6. Spoken fluency (Europe 2001:129)

#### **Interviewer Guidelines**

In this interview, I'd like to find out how useful the CEFR descriptors are for describing the kind of professional English that you use when lecturing here in Finland. I'll be showing you the scales and asking you some questions related to each category in order to find out what you think about different concepts in the scales as well as to find out how well the descriptors cover your spoken English for the purpose of lecturing.

#### **General questions**

1. How familiar are you with the CEFR scales?
2. How well do you feel you can use these CEFR descriptors for self-evaluation?
3. This table shows the six CEFR levels on a global scale. Read through the scale and place yourself for two purposes: a) for professional and b) for non-professional.

#### **Grammatical accuracy scale**

1. Look at this scale for grammatical accuracy. Read through the descriptors.
2. Where would you place yourself on this scale for non-professional purposes?  
For professional purposes?
3. What guided your choice? Can you pinpoint key words or phrases?
4. What do you think would be the minimal level required for lecturing your subjects?
5. Looking at this scale, what do you think is the measure for grammatical accuracy? What does 'error' mean?
6. Is this type of grammatical description applicable to your situation at work?
7. What does grammatical accuracy mean to you? How would you define your own measure of grammatical accuracy?
8. How well do you think these scales cover the kind of professional English that you use in lectures?

### **Vocabulary range scale**

1. Look at this scale for vocabulary range. Read through the descriptors.
2. Where would you place yourself on the scale for non-professional purposes?  
And for professional purposes?
3. What guided your choice? Can you pinpoint key words or phrases?
4. What do you think would be the minimal level required for lecturing your subjects?
5. What do you think about these descriptors as a measure of your vocabulary range for work? Do they measure the kind of vocabulary that you need when lecturing?
6. How about for measuring your vocabulary for everyday use? Any differences from the vocabulary you would need for lecturing?
7. When describing language, what do you understand from terms like 'idiomatic expression' and 'colloquialisms'? What do those terms mean to you?
8. When lecturing, do you use 'idiomatic expressions' and 'colloquialisms'? Why or why not?
9. How well do you think these scales cover the kind of professional vocabulary that you use in lectures?

### **Phonological control scale**

1. In terms of pronunciation and intonation, where would you place yourself on this scale for non-professional purposes? And for professional purposes?
2. What guided your choice? Can you pinpoint key words or phrases?
3. What do you think would be the minimal level required for lecturing your subjects?
4. In this scale, what do you think the measure of pronunciation is?
5. What does the term foreign accent mean to you?
6. How about your own pronunciation target for English, what would you say it is?
7. To be intelligible in lectures, what do you consider to be important? To be intelligible, do you need to speak like a NS?
8. Could you describe what you think would be the ideal pronunciation target for lecturing in English here in Finland?
9. Do you interact more with NSs or NNs?
10. How well do you think these scales cover the kind of pronunciation that you think would be appropriate for your lectures?

### **Coherence and cohesion scale**

1. Are you familiar with the terms coherence and cohesion?
2. Where would you place yourself on this scale for non-professional purposes?  
And for professional purposes?
3. What guided your choice? Can you pinpoint key words or phrases?
4. What do you think would be the minimal level required for lecturing your subjects?
5. How well do you think these scales cover the kind of techniques that you use to achieve coherence and cohesion in lectures?

### **Spoken fluency scale**

1. In terms spoken fluency, where would you place yourself on this scale for non-professional purposes? And for professional purposes?
2. What guided your choice? Can you pinpoint key words or phrases?
3. What do you think would be the minimal level required for lecturing your subjects?
4. How would you describe your fluency target for lecturing?
5. To speak well in lectures, what do you think you need to be able to do?
6. How well do you think these scales cover the kind of spoken fluency that you think would be suitable in your lectures?
7. What do you think about the descriptors at the top of this scale? Would they be applicable to your situation at work – for lecturing? Why or why not?

## Appendix E: Guide for stimulated recall

*These instructions and questions were based on an example in the appendix of Gass and Mackey 2000.*

### Instructions

What we're going to do now is watch the video.

We are interested in what you were thinking at the time you were talking about the lecture contents.

We can see what you were doing by looking at the video, but we don't know what you were thinking.

So what I'd like you to do is tell me what you were thinking, what was in your mind at the time while you were talking to the class.

Let's start the video and you can pause the video any time that you want. So if you want to tell me something about what you were thinking, you can push pause.

If I have a question about what you were thinking, then I'll ask you to push pause and ask you to talk about that part of the video.

### Questions

- What were you thinking just then/at this point?
- Tell me what you were thinking at that point. Reason?
- How did you feel? Reason?
- To what extent did you feel that the students were with you? Reason?
- Did you have any particular objectives in mind in this segment? If so, what were they?

### Post video

- Any thoughts you'd like to share about this video experience?
- I'm curious, what are your impressions of the lecture you gave?
- Anything else you'd like to add?

## Appendix F: Student questionnaire

### Aalto University - Pilot Mentoring Program April 2011

We would like to ask you to help us by answering the following questions concerning your studies in English at Aalto University School of Science. This survey is being conducted as part of the Pilot Mentoring Program to better understand teaching and learning through English in a multicultural environment. This is not a test so there are no "right" or "wrong" answers and you don't even need to write your name on it. We are interested in your personal opinion. Please give your answers sincerely as only this will guarantee the success of investigation. Your participation is greatly appreciated!

**I. The following questions pertain to the lecture you just attended.  
Please circle the alternative corresponding to your opinion.**

	Agree	Somewhat agree	Somewhat disagree	Disagree
1. The lecturer expresses himself/herself well.				
2. I had some difficulty understanding the contents of the lecture. [reversed coding]				
3. The length of the lecturer's sentences are just what I need.				
4. The lecturer makes some grammatical mistakes. [reversed coding]				
5. I would understand the lecture better in my native language. [reversed coding]				
6. The topic of the lecture is very challenging. [reversed coding]				
7. It is easy to follow the lecture.				
8. The lecturer goes through the contents too quickly. [reversed coding]				
9. The atmosphere during the lecture is relaxed.				
10. The contents of the lecture are presented logically.				
11. The lecturer presents clear, smoothly flowing arguments.				
12. I feel that I am learning the key terminology for the course.				
13. Overall, I understand the contents of the lecture well.				
14. The lecturer presents ideas in a way that helps me to notice significant points.				
15. I find it hard to understand the lecturer's pronunciation. [reversed coding]				
16. The lecturer makes disturbing grammatical mistakes. [reversed coding]				
17. Overall, the language used disturbs my concentration on the contents of the lecture.				

[reversed coding]

18. Most of the lecture remained unclear to me.

[reversed coding]

19. The lecturer makes very few grammatical mistakes, if any.

20. The lecturer speaks too fast.

[reversed coding]

21. The lecturer's sentences are too short. [reversed coding]

22. I feel uneasy for the lecturer while s/he lectures in English.

[reversed coding]

23. I would prefer a native speaker of English as a lecturer.

[reversed coding]

24. I understand the key concepts of the lecture.

25. The lecturer presents clear, detailed descriptions of the course contents.

26. I feel that I am learning the key concepts in this course.

27. I find it easy to listen to the lecturer when s/he speaks.

28. I understand most of the ideas presented in the lecture.

29. The lecturer takes contact with the audience.

30. The atmosphere in the lecture encourages discussion and questions.

31. The course terminology is well represented in the lecture.

32. Overall, my impression is that the lecturer presents a clear lecture that I can understand.

## II. The following questions pertain to the lecturer's English skills.

Please circle the alternative(s) corresponding to your opinion.

33. To follow the lecturer's language was

1. Very easy
2. Easy
3. Not so easy
4. Difficult

34. The lecturer's English was challenging because of

1. Vocabulary
2. Fluency
3. Intonation (=melody of speech)
4. Pronouncing single sounds
5. Other, specify

## III. Please add your comments.

35. Further comments:

---

---

---

---

## IV. Please circle the correct alternative or write your answer in the space reserved for it.

36. My native language is                      1 Finnish                      2 Swedish                      3 Other, specify \_\_\_\_\_

37. The level of my English skills in my opinion is                      1 Excellent                      2 Good                      3 Fair                      4 Poor

Please add your student number:

Thank you for your responses and for your help!

## Appendix H: Transcription conventions

The transcription conventions are based on a slightly modified version of the ELFA corpus guide. Special symbols used in the transcripts are explained below. The original guide is available at <http://www.helsinki.fi/englanti/elfa/ELFA%20transcription%20guide.pdf>

### Speaker codes

I	Interviewer
A-G	Lecturer

### Transcription symbols

<S></S>	Utterance begins/ends
,	Brief pause 2-3 sec
.	Pause 3-4 sec
[text]	Overlapping speech (approximate, shown to the nearest word, words not split by overlap tags)
C-A-P	Capital letters for spelling out a word or acronym
(xx)	Unintelligible speech
((...))	Omitted text from transcription
@ @	Laughter
@text@	Spoken laughter
<NAME>	Name of participants
<TEXT>	Descriptions and comments between tags
<READING>text<READING>	Reading aloud

### Errata sheet: print vs e-version

<b>Page</b>	<b>Print version</b>	<b>e-version</b>	<b>Difference</b>
23	H 3	H3	spacing
23	H3	This hypothesis	wording
66 (Figure 4-1 caption)	languaage	language	typo
106 (Table 6)	Arabic/Frenc h	Arabic/French	alignment